

PCT

10869309  
WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> :  
C07K 14/00

A2

(11) International Publication Number: WO 98/40404

(43) International Publication Date: 17 September 1998 (17.09.98)

(21) International Application Number: PCT/US98/04601

(22) International Filing Date: 9 March 1998 (09.03.98)

(30) Priority Data:

08/815,381 11 March 1997 (11.03.97) US  
09/036,321 6 March 1998 (06.03.98) US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

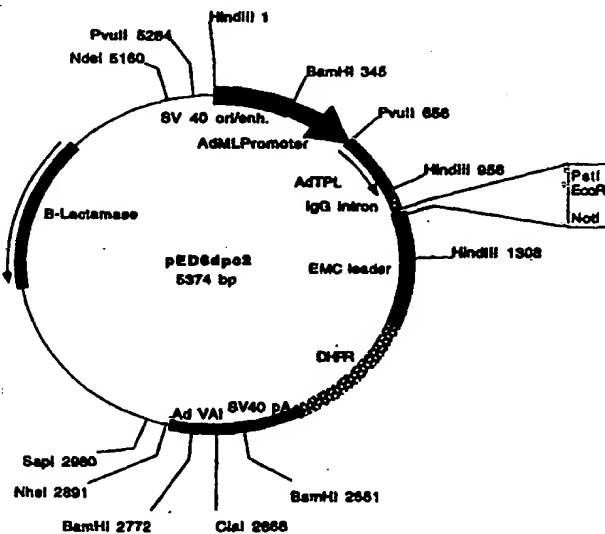
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2  
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. GST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5

## SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

10 This application is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/815,381), filed March 11, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

15 The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

20 Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein

25 in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of

30 DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUBSTITUTE SHEET (RULE 26)

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID  
5 NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID  
NO:1 from nucleotide 521 to nucleotide 1111;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID  
NO:1 from nucleotide 536 to nucleotide 817;

10 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ax318\_3 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the  
cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353;

15 (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ax318\_3 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the  
cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353;

20 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(i) a polynucleotide encoding a protein comprising a fragment of the  
amino acid sequence of SEQ ID NO:2 having biological activity, the fragment  
comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ  
25 ID NO:2;

(j) a polynucleotide which is an allelic variant of a polynucleotide of  
(a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein  
of (h) or (i) above ; and

30 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 521 to nucleotide 1111; the nucleotide sequence of SEQ ID NO:1 from nucleotide 536 to nucleotide 817; the nucleotide sequence of the full-length protein



coding sequence of clone ax318\_3 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone ax318\_3 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353. In yet other preferred  
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2 from amino acid 4 to amino acid 99.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

10 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:2;
  - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 4 to  
15 amino acid 99;
  - (c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:2; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone  
20 ax318\_3 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2 or the amino acid sequence of SEQ ID NO:2 from amino acid 4 to amino acid 99.

In one embodiment, the present invention provides a composition comprising an  
25 isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 61 to nucleotide 864;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID  
30 NO:3 from nucleotide 826 to nucleotide 1386;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 61 to nucleotide 864; the nucleotide sequence of SEQ ID NO:3 from nucleotide 826 to nucleotide 1386; the nucleotide sequence of the full-length protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:31 from amino acid 148 to amino acid 249.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
- 5 (b) the amino acid sequence of SEQ ID NO:31;
- (c) the amino acid sequence of SEQ ID NO:31 from amino acid 148 to amino acid 249;
- (d) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:4;
- 10 (e) fragments of the amino acid sequence of SEQ ID NO:31 comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:31; and
- (f) the amino acid sequence encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;
- 15 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4, the amino acid sequence of SEQ ID NO:31, or the amino acid sequence of SEQ ID NO:31 from amino acid 148 to amino acid 249.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 77 to nucleotide 1624;
- 25 (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 390 to nucleotide 789;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bg465\_2 deposited under accession number ATCC 98353;
- 30 (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bg465\_2 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

5 (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 253 to amino acid 262 of SEQ ID NO:6;

10 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

15 Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 77 to nucleotide 1624; the nucleotide sequence of SEQ ID NO:5 from nucleotide 390 to nucleotide 789; the nucleotide sequence of the full-length protein coding sequence of clone bg465\_2 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone bg465\_2 deposited  
20 under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 260 to amino acid  
25 343.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group  
30 consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 260 to amino acid 343;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 253 to amino acid 262 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353; the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 260 to amino acid 343.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 48 to nucleotide 1055;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 216 to nucleotide 1055;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 494 to nucleotide 958;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk291\_3 deposited under accession number ATCC 98353;

(g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bk291\_3 deposited under accession number ATCC 98353;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:8;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

5 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 48 to nucleotide 1055; the nucleotide sequence of SEQ ID NO:7 from nucleotide 216 to nucleotide 1055; the nucleotide sequence of SEQ ID NO:7 from  
10 nucleotide 494 to nucleotide 958; the nucleotide sequence of the full-length protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of  
15 clone bk291\_3 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 188 to amino acid 306.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ  
20 ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;  
25 (b) the amino acid sequence of SEQ ID NO:8 from amino acid 188 to amino acid 306;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone  
30 bk291\_3 deposited under accession number ATCC 98353;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 188 to amino acid 306.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

5 (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 64 to nucleotide 1197;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1 to nucleotide 828;

10 (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353;

15 (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353;

20 (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 184 to amino acid 193 of SEQ ID NO:10;

25 (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

30 (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 64 to nucleotide 1197; the nucleotide sequence of SEQ ID NO:9 from nucleotide 1 to nucleotide 828; the nucleotide sequence of the full-length protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353; or the

nucleotide sequence of the mature protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353. In yet other preferred  
5 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 255.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

10 In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:10;
- (b) the amino acid sequence of SEQ ID NO:10 from amino acid 1 to  
15 amino acid 255;

- (c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 184 to amino acid 193 of SEQ ID NO:10; and

- (d) the amino acid sequence encoded by the cDNA insert of clone  
20 bp537\_4 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 255.

In one embodiment, the present invention provides a composition comprising an  
25 isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 581 to nucleotide 1534;

- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 928 to nucleotide 1510;

- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353;



(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 154 to amino acid 163 of SEQ ID NO:12;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 581 to nucleotide 1534; the nucleotide sequence of SEQ ID NO:11 from nucleotide 928 to nucleotide 1510; the nucleotide sequence of the full-length protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 150 to amino acid 310.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:12;
- 5 (b) the amino acid sequence of SEQ ID NO:12 from amino acid 150 to amino acid 310;
- (c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 154 to amino acid 163 of SEQ ID NO:12; and
- 10 (d) the amino acid sequence encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 150 to amino acid 310.

- 15 In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID
- 20 NO:13 from nucleotide 29 to nucleotide 2227;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1334 to nucleotide 2227;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 746;
- 25 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;
- 30 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 361 to amino acid 370 of SEQ ID NO:14;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 29 to nucleotide 2227; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1334 to nucleotide 2227; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 746; the nucleotide sequence of the full-length protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 239, or a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14 from amino acid 119 to amino acid 733.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 239;

(c) the amino acid sequence of SEQ ID NO:14 from amino acid 119 to amino acid 733;

(d) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 361 to amino acid 370 of SEQ ID NO:14;  
5 and

(e) the amino acid sequence encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14, the amino acid sequence of  
10 SEQ ID NO:14 from amino acid 1 to amino acid 239, or the amino acid sequence of SEQ ID NO:14 from amino acid 119 to amino acid 733.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID  
15 NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 364 to nucleotide 777;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 636;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment

comprising the amino acid sequence from amino acid 64 to amino acid 73 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 364 to nucleotide 777; the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 636; the nucleotide sequence of the full-length protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 91.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 91;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 64 to amino acid 73 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 91.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 619 to nucleotide 1434;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 520 to nucleotide 1323;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:18;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:17 from nucleotide 619 to nucleotide 1434; the nucleotide sequence of SEQ ID NO:17 from nucleotide 520 to nucleotide 1323; the nucleotide sequence of the full-length protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353; or the  
5 nucleotide sequence of the mature protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353. In yet other preferred  
10 embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 235.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:17.

In other embodiments, the present invention provides a composition comprising  
15 a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:18;
- (b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 235;
- 20 (c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:18; and
- (d) the amino acid sequence encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;
- 25 the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:18 or the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 235.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

- 30 (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2063 to nucleotide 2290;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2276 to nucleotide 2290;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2037 to nucleotide 2405;

5 (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;

10 (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;

15 (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 33 to amino acid 42 of SEQ ID NO:20;

20 (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

25 (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:19 from nucleotide 2063 to nucleotide 2290; the nucleotide sequence of SEQ ID NO:19 from nucleotide 2276 to nucleotide 2290; the nucleotide sequence of SEQ ID NO:19 from nucleotide 2037 to nucleotide 2405; the nucleotide sequence of the full-length protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353; or the nucleotide sequence of the mature protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353. In other preferred embodiments, the polynucleotide encodes the full-length or mature protein encoded by the cDNA insert of



clone ep234\_1 deposited under accession number ATCC 98353. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 69.

5 Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:19.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

- 10 (a) the amino acid sequence of SEQ ID NO:20;  
(b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 69;  
(c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 33 to amino acid 42 of SEQ ID NO:20;  
15 and  
(d) the amino acid sequence encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:20 or the amino acid sequence  
20 of SEQ ID NO:20 from amino acid 1 to amino acid 69.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or  
25 modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

- (a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and  
30 (b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

#### DETAILED DESCRIPTION

##### ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

##### Clone "ax318\_3"

A polynucleotide of the present invention has been identified as clone "ax318\_3". ax318\_3 was isolated from a human adult testes cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ax318\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ax318\_3 protein").

The nucleotide sequence of ax318\_3 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ax318\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ax318\_3 should be approximately 1200 bp.

The nucleotide sequence disclosed herein for ax318\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ax318\_3 demonstrated at least some similarity with sequences identified as AA255872 (zs19a01.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone), AA625610 (zv89h04.s1 Soares NhHMPu S1 Homo sapiens cDNA clone 766999 3' similar to WP:K10B2.1 CE02008 TRANSDUCIN BETA CHAIN), D86043 (Human mRNA for SHPS-1, complete cds), H23217 (ym52f03.s1 Homo sapiens cDNA clone 51880 3'), U06701 (Human clone CCA53 mRNA containing CCA trinucleotide repeat), and W05822 (za90b02.r1 Soares fetal lung NbHL19W Homo sapiens cDNA clone 299787 5'). The predicted amino acid sequence disclosed herein for ax318\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted ax318\_3 protein demonstrated at least some similarity to sequences identified as D85183 (SHPS-1 [Rattus rattus]), D86043 (SHPS-1 [Homo sapiens]), R85852 (WD-40 domain-containing beta-transducin protein), Y10376 (SIRP-beta1 [Homo sapiens]), and Z79757 (F55B12.3 [Caenorhabditis elegans]). SHPS-1 is a membrane glycoprotein that binds the SH2-domain-containing protein tyrosine phosphatase SHP-2 in response to mitogens and cell adhesion (Fujioka *et al.*, 1996, *Mol. Cell. Biol.* 16(12): 6887-6899, incorporated by reference herein). Based upon sequence similarity, ax318\_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a putative transmembrane domain within the ax318\_3 protein sequence, centered around amino acid 28 of SEQ ID NO:2; amino acids 15 to 27 of SEQ ID NO:2 are

also a possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 28.

Clone "bg140\_1"

5 A polynucleotide of the present invention has been identified as clone "bg140\_1". bg140\_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bg140\_1 is a full-length clone,  
10 including the entire coding sequence of a secreted protein (also referred to herein as "bg140\_1 protein").

The nucleotide sequence of bg140\_1 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bg140\_1 protein corresponding to the foregoing  
15 nucleotide sequence is reported in SEQ ID NO:4. Another possible bg140\_1 reading frame and predicted amino acid sequence, encoded by base pairs 641 to 1648 of SEQ ID NO:3, is reported in SEQ ID NO:31. A frameshift in the nucleotide sequence of SEQ ID NO:3, caused, for example, by a single base pair deletion near position 825 of SEQ ID NO:3, could join the open reading frame encoding SEQ ID NO:4 with that encoding SEQ ID  
20 NO:31.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bg140\_1 should be approximately 2900 bp.

The nucleotide sequence disclosed herein for bg140\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and  
25 FASTA search protocols. bg140\_1 demonstrated at least some similarity with sequences identified as AA075629 (zm89a04.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA clone 545070 3'), AA113989 (zn27h12.r1 Stratagene neuroepithelium NT2RAMI 937234 Homo sapiens cDNA clone 548711 5'), N77135 (yz85h11.r1 Homo sapiens cDNA clone 289893 5'), R11701 (yf49d02.r1 Homo sapiens cDNA clone 25600 5'), R13178  
30 (yf73e04.r1 Homo sapiens cDNA clone 27855 5'), and Y14946 (Homo sapiens mRNA for SPIN protein). The predicted amino acid sequence disclosed herein for bg140\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bg140\_1 protein demonstrated at least some similarity to sequences identified as AF038969 (general transcription factor 2-I [Homo

sapiens]), U77948 (Bruton's tyrosine kinase-associated protein-135; BAP-135 [Homo sapiens]), and Y14946 (SPIN protein [Homo sapiens]). Based upon sequence similarity, bg140\_1 proteins and each similar protein or peptide may share at least some activity.

5        Clone "bg465\_2"

A polynucleotide of the present invention has been identified as clone "bg465\_2". bg465\_2 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer  
10 analysis of the amino acid sequence of the encoded protein. bg465\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bg465\_2 protein").

The nucleotide sequence of bg465\_2 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the  
15 predicted amino acid sequence of the bg465\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bg465\_2 should be approximately 2500 bp.

The nucleotide sequence disclosed herein for bg465\_2 was searched against the  
20 GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bg465\_2 demonstrated at least some similarity with sequences identified as AA133150 (zm25b07.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 526645 5' similar to WP:F07F6.4 CE01896 ZINC FINGER PROTEIN), AA144842 (mr69h01.r1 Stratagene mouse testis (#937308) Mus musculus cDNA clone 602737 5'),  
25 AA484561 (nf06a12.s1 NCI\_CGAP\_Li1 Homo sapiens cDNA clone), D16939 (Human HepG2 3' region cDNA, clone hmd5d06), L24125 (Saccharomyces cerevisiae zinc finger protein (GCS1) gene, complete cds), R18422 (yg02f06.r1 Homo sapiens cDNA clone 30950 5'), T32543 (EST50558 Homo sapiens cDNA 5' end similar to None), W88569 (zh70b12.s1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 417407 3'), Z74274  
30 (S.cerevisiae chromosome IV reading frame ORF YDL226c), and Z82199 (Human DNA sequence \*\*\* SEQUENCING IN PROGRESS \*\*\* from clone 316D5; HTGS phase 1). The predicted amino acid sequence disclosed herein for bg465\_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bg465\_2 protein demonstrated at least some similarity to sequences

identified as U23486 (similar to *S. cerevisiae* zinc finger protein GCS1 (SP GCS1\_YEAST) [*Caenorhabditis elegans*]). Based upon sequence similarity, bg465\_2 proteins and each similar protein or peptide may share at least some activity.

5        Clone "bk291\_3"

A polynucleotide of the present invention has been identified as clone "bk291\_3". bk291\_3 was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer  
10       analysis of the amino acid sequence of the encoded protein. bk291\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bk291\_3 protein").

The nucleotide sequence of bk291\_3 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the  
15       predicted amino acid sequence of the bk291\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Amino acids 44 to 56 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 57, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone  
20       bk291\_3 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for bk291\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bk291\_3 demonstrated at least some similarity with sequences identified as AA135915 (zl19f04.r1 Soares pregnant uterus NbHPU Homo sapiens cDNA  
25       clone 502399 5'), AA156738 (zl19f04.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 502399 3' similar to WP:T05E11.5 CE06364 YEAST YKK0 LIKE), AF017785 (Mus musculus MHC Class I H13a minor histocompatibility peptide (H13) mRNA, partial cds), N40134 (yw73a12.r1 Homo sapiens cDNA clone 257854 5' similar to SW:YKK0\_YEAST P34248 HYPOTHETICAL 67.5 KD PROTEIN IN APE1/LAP4-CWP1  
30       INTERGENIC REGION), N42726 (yy11d01.r1 Homo sapiens cDNA clone 270913 5' similar to SW:YKK0\_YEAST P34248 HYPOTHETICAL 67.5 KD PROTEIN IN APE1/LAP4-CWP1 INTERGENIC REGION), R67144 (yi31h04.r1 Homo sapiens cDNA clone 140887 5' similar to SP:YKK0\_YEAST P34248 HYPOTHETICAL 67.5 KD PROTEIN IN APE1/LAP4-MBR1 INTERGENIC), R78822 (yi90b05.r1 Homo sapiens cDNA clone 146481 5'), R79317

(yi90b05.s1 Homo sapiens cDNA clone 146481 3'), T23944 (Human gene signature HUMGS05889), and W04243 (za58h10.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone). The predicted amino acid sequence disclosed herein for bk291\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted bk291\_3 protein demonstrated at least some similarity to sequences identified as X71133 (YKL450 [Saccharomyces cerevisiae]), Z28100 (ORF YKL100c [Saccharomyces cerevisiae]), and Z68751 (T05E11.5 [Caenorhabditis elegans]). Based upon sequence similarity, bk291\_3 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts six additional potential transmembrane domains within the bk291\_3 protein sequence, centered around amino acids 70, 125, 170, 220, 260, and 280 of SEQ ID NO:8.

Clone "bp537\_4"

A polynucleotide of the present invention has been identified as clone "bp537\_4". bp537\_4 was isolated from a human fetal kidney cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. bp537\_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "bp537\_4 protein").

The nucleotide sequence of bp537\_4 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the bp537\_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone bp537\_4 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for bp537\_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. bp537\_4 demonstrated at least some similarity with sequences identified as AA007576 (zh99b06.r1 Soares fetal liver spleen 1NFLS S1 Homo sapiens cDNA clone 429395 5'), AA287563 (zs52g02.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA 5'), R09093 (yf21h09.r1 Homo sapiens cDNA clone 127553 5'), and T33088 (EST56631 Homo sapiens cDNA 5' end similar to None). The predicted amino acid sequence disclosed herein for bp537\_4 was searched against the GenPept and GeneSeq amino acid

sequence databases using the BLASTX search protocol. The predicted bp537\_4 protein demonstrated at least some similarity to sequences identified as U34998 (Rad9 [Coprinus cinereus]). Based upon sequence similarity, bp537\_4 proteins and each similar protein or peptide may share at least some activity.

5

Clone "cs431\_2"

A polynucleotide of the present invention has been identified as clone "cs431\_2". cs431\_2 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
10 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cs431\_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cs431\_2 protein").

The nucleotide sequence of cs431\_2 as presently determined is reported in SEQ ID  
15 NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cs431\_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cs431\_2 should be approximately 2300 bp.

20 The nucleotide sequence disclosed herein for cs431\_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cs431\_2 demonstrated at least some similarity with sequences identified as AA021033 (ze67h12.s1 Soares retina N2b4HR Homo sapiens cDNA clone 364103 3' similar to contains Alu repetitive element), H29349 (ym32c02.r1 Homo sapiens  
25 cDNA clone 49839 5' similar to SP:KYES\_XIPHE P27447 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE YES), L14577 (Homo sapiens cystathionine beta-synthase (CBS) mRNA, complete cds), Q87430 (Human cystathionine beta-synthase cDNA), and X92659 (H.sapiens intron 4 from p53 gene). The predicted amino acid sequence disclosed herein for cs431\_2 was searched against the GenPept and GeneSeq amino acid sequence  
30 databases using the BLASTX search protocol. The predicted cs431\_2 protein demonstrated at least some similarity to sequences identified as L19501 (cystathionine beta-synthase [Homo sapiens]), R71376 (Human cystathionine beta-synthase), U61167 (SH3 domain-containing protein SH3P18 [Homo sapiens]), and X82166 (cystathionine beta-synthase [Homo sapiens]). Based upon sequence similarity, cs431\_2 proteins and



each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the cs431\_2 protein sequence centered around amino acid 30 of SEQ ID NO:12. The nucleotide sequence of cs431\_2 indicates that it may contain an Alu repetitive element.

5

Clone "cw976\_1"

A polynucleotide of the present invention has been identified as clone "cw976\_1". cw976\_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
10 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cw976\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cw976\_1 protein").

The nucleotide sequence of cw976\_1 as presently determined is reported in SEQ  
15 ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cw976\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14. Amino acids 423 to 435 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 436, or are a transmembrane domain.

20 The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cw976\_1 should be approximately 3300 bp.

The nucleotide sequence disclosed herein for cw976\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cw976\_1 demonstrated at least some similarity with sequences  
25 identified as AA280747 (zs96a09.r1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:711448 5'), AA323946 (EST26798 Cerebellum II Homo sapiens cDNA 5' end), R13665 (yf60g06.r1 Homo sapiens cDNA clone 26764 5'), R16892 (yf44d06.s2 Homo sapiens cDNA clone 129707 3'), R85177 (yo43d07.r1 Homo sapiens cDNA clone 180685 5'), and R89648 (ym97c02.r1 Homo sapiens cDNA clone 166850 5'). Based upon sequence  
30 similarity, cw976\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the cw976\_1 protein sequence at the amino terminus of SEQ ID NO:14.

Clone "cw1233\_3"

A polynucleotide of the present invention has been identified as clone "cw1233\_3". cw1233\_3 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was  
5 identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. cw1233\_3 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "cw1233\_3 protein").

The nucleotide sequence of cw1233\_3 as presently determined is reported in SEQ  
10 ID NO:15. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the cw1233\_3 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone cw1233\_3 should be approximately 3000 bp.

15 The nucleotide sequence disclosed herein for cw1233\_3 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. cw1233\_3 demonstrated at least some similarity with sequences identified as R61600 (yh16f08.r1 Homo sapiens cDNA clone 37807 5'), R83929 (yp06h11.s1 Homo sapiens cDNA clone 186693 3' similar to contains Alu repetitive element; contains  
20 TAR1 repetitive element), R87877 (yo45h05.r1 Homo sapiens cDNA clone 180921 5'), U14568 (\*\*\*ALU WARNING Human Alu-Sb subfamily consensus sequence). The predicted amino acid sequence disclosed herein for cw1233\_3 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted cw1233\_3 protein demonstrated at least some similarity to sequences  
25 identified as AB002317 (KIAA0319 [Homo sapiens]). Based upon sequence similarity, cw1233\_3 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of cw1233\_3 indicates that it may contain an Alu repetitive element.

Clone "dg1\_1"

30 A polynucleotide of the present invention has been identified as clone "dg1\_1". dg1\_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer

analysis of the amino acid sequence of the encoded protein. dg1\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "dg1\_1 protein").

The nucleotide sequence of dg1\_1 as presently determined is reported in SEQ ID NO:17. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the dg1\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:18.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone dg1\_1 should be approximately 1900 bp.

The nucleotide sequence disclosed herein for dg1\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. dg1\_1 demonstrated at least some similarity with sequences identified as AA731281 (nw57f05.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE 1250721 similar to gb Y00345\_cds1 POLYADENYLATE-BINDING PROTEIN (HUMAN)). The predicted amino acid sequence disclosed herein for dg1\_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted dg1\_1 protein demonstrated at least some similarity to sequences identified as M81878 (hyaluronidase [Clostridium perfringens]) and U28742 (similar to hyaluronoglucosaminidase (SPNAGH\_CLOPE, P26831) [Caenorhabditis elegans]). Hyaluronoglucosaminidase is a secreted protein found, for example, in bee venom. Based upon sequence similarity, dg1\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the dg1\_1 protein sequence centered around amino acid 150 of SEQ ID NO:18.

#### Clone "ep234\_1"

A polynucleotide of the present invention has been identified as clone "ep234\_1". ep234\_1 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ep234\_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ep234\_1 protein").

The nucleotide sequence of ep234\_1 as presently determined is reported in SEQ ID NO:19. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ep234\_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:20. Amino acids 59 to 71 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 72, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ep234\_1 should be approximately 2500 bp.

The nucleotide sequence disclosed herein for ep234\_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ep234\_1 demonstrated at least some similarity with sequences identified as AA044714 (zf54d09.r1 Soares retina N2b4HR Homo sapiens cDNA clone 380753 5'), AA454840 (zx79d09.s1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 809969 3'), AA470137 (zu11e01.s1 Soares testis NHT Homo sapiens cDNA clone 731544 3'), H24852 (yl42c11.r1 Homo sapiens cDNA clone 160916 5'), N64648 (yz87b06.s1 Homo sapiens cDNA clone 290003 3'), R45788 (Ha662-f Homo sapiens cDNA clone a662-f), T70546 (yd15b07.s1 Homo sapiens cDNA clone 108277 3'), W73481 (zd54e01.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 344472 3'), and W73553 (zd54e01.r1 Soares fetal heart). Based upon sequence similarity, ep234\_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts at least two potential transmembrane domains within the ep234\_1 protein sequence of SEQ ID NO:20.

#### Deposit of Clones

Clones ax318\_3, bg140\_1, bg465\_2, bk291\_3, bp537\_4, cs431\_2, cw976\_1, cw1233\_3, dg1\_1, and ep234\_1 were deposited on March 11, 1997 with the American Type Culture Collection as an original deposit under the Budapest Treaty and were given the accession number ATCC 98353, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b).

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the

appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Fig. 1. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of the oligonucleotide probe that was used to isolate each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

<u>Clone</u>	<u>Probe Sequence</u>
ax318_3	SEQ ID NO:21
bg140_1	SEQ ID NO:22
bg465_2	SEQ ID NO:23, SEQ ID NO:32
bk291_3	SEQ ID NO:24
bp537_4	SEQ ID NO:25
cs431_2	SEQ ID NO:26
cw976_1	SEQ ID NO:27
cw1233_3	SEQ ID NO:28
dg1_1	SEQ ID NO:29
ep234_1	SEQ ID NO:30

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoramidite residue rather than a

nucleotide (such as, for example, that produced by use of biotinyl phosphoramidite (1-dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

- (a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;
  - (b) It should be designed to have a  $T_m$  of approx.  $80^\circ\text{C}$  (assuming  $2^\circ$  for each A or T and 4 degrees for each G or C).
- 10 The oligonucleotide should preferably be labeled with  $g\text{-}^{32}\text{P}$  ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated
- 15 by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately  $4\text{e}+6$  dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100  $\mu\text{l}$  of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100  $\mu\text{g}/\text{ml}$ . The culture should preferably be

20 grown to saturation at  $37^\circ\text{C}$ , and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100  $\mu\text{g}/\text{ml}$  and agar at 1.5% in a 150 mm petri dish when grown overnight at  $37^\circ\text{C}$ . Other

25 known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at  $65^\circ\text{C}$  for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100  $\mu\text{g}/\text{ml}$  of yeast RNA, and 10 mM EDTA (approximately

30 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to  $1\text{e}+6$  dpm/mL. The filter is then preferably incubated at  $65^\circ\text{C}$  with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed

by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes. A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, *Bio/Technology* 10, 773-778 (1992) and in R.S. McDowell, *et al.*, *J. Amer. Chem. Soc.* 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can

be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that  
5 has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense  
10 polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* 15(7): 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* 62(1): 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* 58: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed  
15 herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein).  
20 In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of  
25 transposable elements (Plasterk, 1992, *Bioessays* 14(9): 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* 90(16): 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* 91(2): 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* 336: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614,396;  
30 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of



assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25% (more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide, as determined by those of skill in the art. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologs are those isolated from mammalian species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous, or related to that encoded by the polynucleotides.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides capable of hybridizing under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp) <sup>‡</sup>	Hybridization Temperature and Buffer <sup>†</sup>	Wash Temperature and Buffer <sup>†</sup>
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T <sub>B</sub> *; 1xSSC	T <sub>B</sub> *; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T <sub>D</sub> *; 1xSSC	T <sub>D</sub> *; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T <sub>F</sub> *; 1xSSC	T <sub>F</sub> *; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T <sub>H</sub> *; 4xSSC	T <sub>H</sub> *; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T <sub>J</sub> *; 4xSSC	T <sub>J</sub> *; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T <sub>L</sub> *; 2xSSC	T <sub>L</sub> *; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T <sub>N</sub> *; 6xSSC	T <sub>N</sub> *; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T <sub>P</sub> *; 6xSSC	T <sub>P</sub> *; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T <sub>R</sub> *; 4xSSC	T <sub>R</sub> *; 4xSSC

<sup>‡</sup>: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed

to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

\*: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH<sub>2</sub>PO<sub>4</sub>, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

\*T<sub>B</sub> - T<sub>R</sub>: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T<sub>m</sub>) of the hybrid, where T<sub>m</sub> is determined according to the following equations. For hybrids less than 18 base pairs in length, T<sub>m</sub>(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T<sub>m</sub>(°C) = 81.5 + 16.6(log<sub>10</sub>[Na<sup>+</sup>]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na<sup>+</sup>] is the concentration of sodium ions in the hybridization buffer ([Na<sup>+</sup>] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are

provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205

cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli*, *Bacillus subtilis*, *Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as

those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art

(see, e.g., U.S. Patent No. 5,18,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

### USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

#### Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that

described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

5 The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

20 Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

#### 25 Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors  
5 discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3,  
10 MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those  
15 described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology  
20 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a.  
25 Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon  $\gamma$ , Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine  
30 Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in*



*Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; 5 Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: 10 *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., 15 *Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger et al., *Eur. J. Immun.* 11:405-411, 1981; Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988.

#### Immune Stimulating or Suppressing Activity:

20 A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well 25 as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, 30 herpesviruses, mycobacteria, *Leishmania* spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the

molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to  
5 anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

10 The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as  
15 described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

20 Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms.

25 Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from  
30 the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythematosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and

murine experimental myasthenia gravis (see Paul ed., *Fundamentals of Immunology*, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy.

5 Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B  
10 lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune  
15 responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.  
20

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present  
25 invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The  
30 transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary

costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I  $\alpha$  chain protein and  $\beta_2$  microglobulin protein or an MHC class II  $\alpha$  chain protein and an MHC class II  $\beta$  chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowman et al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro*

antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: *Current Protocols in Immunology*, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., *J. Immunol.* 137:3494-3500, 1986; Takai et al., *J. Immunol.* 140:508-512, 1988; Bertagnolli et al., *J. Immunol.* 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., *J. Immunol.* 134:536-544, 1995; Inaba et al., *Journal of Experimental Medicine* 173:549-559, 1991; Macatonia et al., *Journal of Immunology* 154:5071-5079, 1995; Porgador et al., *Journal of Experimental Medicine* 182:255-260, 1995; Nair et al., *Journal of Virology* 67:4062-4069, 1993; Huang et al., *Science* 264:961-965, 1994; Macatonia et al., *Journal of Experimental Medicine* 169:1255-1264, 1989; Bhardwaj et al., *Journal of Clinical Investigation* 94:797-807, 1994; and Inaba et al., *Journal of Experimental Medicine* 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., *Cytometry* 13:795-808, 1992; Gorczyca et al., *Leukemia* 7:659-670, 1993; Gorczyca et al., *Cancer Research* 53:1945-1951, 1993; Itoh et al., *Cell* 66:233-243, 1991; Zacharchuk, *Journal of Immunology* 145:4037-4045, 1990; Zamai et al., *Cytometry* 14:891-897, 1993; Gorczyca et al., *International Journal of Oncology* 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., *Blood* 84:111-117, 1994; Fine et al., *Cellular Immunology* 155:111-122, 1994; Galy et al., *Blood* 85:2770-2778, 1995; Toki et al., *Proc. Nat. Acad. Sci. USA* 88:7548-7551, 1991.

#### Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell

lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and

Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., *Experimental Hematology* 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, NY. 1994; Long  
5 term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland, H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

10

#### Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns,  
15 incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as  
20 well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal  
25 disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue  
30 destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in



circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

#### Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin  $\alpha$  family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-

$\beta$  group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

#### Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion

- include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

#### Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity.

- 10 As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting  
15 therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

- Assay for hemostatic and thrombolytic activity include, without limitation, those  
20 described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

#### Receptor/Ligand Activity

- 25 A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation,  
30 cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without

limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

- 5        Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 10 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

#### Anti-Inflammatory Activity

- 15        Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or 20 suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin 25 lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

#### Cadherin/Tumor Invasion Suppressor Activity

30        Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human

diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from

forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

- 5 Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present
- 10 invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

- Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995;
- 15 Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

#### Tumor Inhibition Activity

- In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities.
- 20 A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating
- 25 or inhibiting factors, agents or cell types which promote tumor growth.

#### Other Activities

- A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious
- 30 agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms;

effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

## 15 ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects



of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

5 A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

10 The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be  
15 supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

20 The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation,  
25 monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

30 As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active

ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein

of the present invention; and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01  $\mu$ g to about 100 mg (preferably about 0.1mg to about 10 mg, more preferably about 0.1  $\mu$ g to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such

antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in  
5 R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where  
10 abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage,  
15 tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or  
20 tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the  
25 composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

30 The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and

polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropylmethylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF- $\alpha$  and TGF- $\beta$ ), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Jacobs, Kenneth  
McCoy, John M.  
LaVallie, Edward R.  
Racie, Lisa A.  
Merberg, David  
Treacy, Maurice  
Spaulding, Vikki  
Agostino, Michael
- (ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 32
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Genetics Institute, Inc.
  - (B) STREET: 87 CambridgePark Drive
  - (C) CITY: Cambridge
  - (D) STATE: MA
  - (E) COUNTRY: U.S.A.
  - (F) ZIP: 02140
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Sprunger, Suzanne A.
  - (B) REGISTRATION NUMBER: 41,323
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (617) 498-8284
  - (B) TELEFAX: (617) 876-5851

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1328 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGTTAACAAG TGGAATGGAA CTCAAAGACA ATATTCTTGT CTCTGGGAAT GCAGATTCTA	60
CAGTTAAAAT CTGGGATATC AAAACAGGAC AGTGTTTACA AACATTGCAA GGTCCCAACA	120
AGCATCAGAG TGCTGTGACC TGTTTACAGT TCAACAAGAA CTTTGTAATT ACCAGCTCAG	180
ATGATGGAAC TGTAAAACTA TGGGACTTGA AAACGGGTGA ATTTATTCGA AACCTAGTCA	240
CATTGGAGAG TGGGGGGAGT GGGGGAGTTG TGTGGCGGAT CACAGCCTCA AACACAAAGC	300
TGGTGTGTGC AGTTGGGAGT CGGAATGGGA CTGAAGAAAC CAAGCTGCTG GTGCTGGACT	360
TTGATGTGGA CATGAAGTGA AGAGCAGAAA AGATGAATTT GTCCAATTGT GTAGACGATA	420
TACTCCCTGC CCTTCCCCCT GCAAAAAGAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	480
AAAAAAAAAA AAAAAAGCGC AAGCTCCAGG TTTCACCACA ATGCCCATCC CTGCCTCCCC	540
ACTCCACCCA CCTCTGCCTT CCTTACTGCT GTATCTGCTG CTTGAACTGG CAGGAGTCAC	600
ACATGTGTTT CATGTGCAAC AAACGGAGAT GTCACAGACT GTATCAACTG GGGAGTCAAT	660
CATCTTGAGT TGCAGCGTAC CCGATACCTT ACCAAATGGA CCTGTCTTGT GGTTCAGGG	720
AACAGGGCCA AACCGGAAAT TAATCTACAA TTTCAAACAA GGTAAC TTTC CCAGAGTAAA	780
AGAGATTGGA GACACCACCA AGCCTGGCAA CACAGACTTT TCCACCCGCA TCCGTGAAAT	840
CTCTCTTGCT GATGCTGGCA CCTATTACTG CGTGAAGTTC ATAAAAGGAA GAGCTATCAA	900
GGAGTACCAA TCAGGTCGGG GCACTCAGGT GTTTGTTACT GAGCAGAATC CAAGACCTCC	960
CAAGAACAGA CCTGCAGGCA GAGCAGGCTC CAGGGCCCAC CATGATGCCC ATACCTGCCT	1020
CTCGGCCCTG CCTGAGAGAA ACAGCACAAA CTATTTCTGTC CAACCCTGCT GCTGCCTCCG	1080
GCTGCTGGGA CTCACAGGCT TGCTGTCAAA ATAATCCAAA CAGGGAAGGA ACGTACAAGT	1140
AAATAACAAA AGCCCCATA CTCTTCTGAC TCCCTGGAGA CAGCTACTTT TTAGGAGTTT	1200
CATTTGCCTT CTTCAAGAGA GCTTTCTTCC ACTGACATAA AATGCCAGCT TGATCGTACA	1260
ATAAATCTGT CTATTTACCT GGTCCAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1320
AAAAAAAAA	1328

(2) INFORMATION FOR SEQ ID NO:2:



## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 197 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Pro Ile Pro Ala Ser Pro Leu His Pro Pro Leu Pro Ser Leu Leu
1           5           10           15
Leu Tyr Leu Leu Leu Glu Leu Ala Gly Val Thr His Val Phe His Val
20           25           30
Gln Gln Thr Glu Met Ser Gln Thr Val Ser Thr Gly Glu Ser Ile Ile
35           40           45
Leu Ser Cys Ser Val Pro Asp Thr Leu Pro Asn Gly Pro Val Leu Trp
50           55           60
Phe Lys Gly Thr Gly Pro Asn Arg Lys Leu Ile Tyr Asn Phe Lys Gln
65           70           75           80
Gly Asn Phe Pro Arg Val Lys Glu Ile Gly Asp Thr Thr Lys Pro Gly
85           90           95
Asn Thr Asp Phe Ser Thr Arg Ile Arg Glu Ile Ser Leu Ala Asp Ala
100          105          110
Gly Thr Tyr Tyr Cys Val Lys Phe Ile Lys Gly Arg Ala Ile Lys Glu
115          120          125
Tyr Gln Ser Gly Arg Gly Thr Gln Val Phe Val Thr Glu Gln Asn Pro
130          135          140
Arg Pro Pro Lys Asn Arg Pro Ala Gly Arg Ala Gly Ser Arg Ala His
145          150          155          160
His Asp Ala His Thr Cys Leu Ser Ala Leu Pro Glu Arg Asn Ser Thr
165          170          175
Asn Tyr Phe Val Gln Pro Cys Cys Cys Leu Arg Leu Leu Gly Leu Thr
180          185          190
Gly Leu Leu Ser Lys
195

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2835 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCTTGATAAG CCAGCTTCAG GAGTAAAGGA AGAATGGTAT GCCAGAATCA CTAAATTAAG	60
ATGGTGGATC AGCTTTTCTG CAAAAAATTT GCGGAAGCCT TGGGGAGCAC TGAAGCCAAG	120
GCTGTACCGT ACCAAAAATT TGAGGCACAC CCGAATGATC TGTACGTGGA AGGACTGCCA	180
GAAAACATTC CTTTCCGAAG TCCCTCATGG TATGGAATCC CAAGGCTGGA AAAAATCATT	240
CAAGTGGGCA ATCGAATTAA ATTTGTTATT AAAAGACCAG AACTTCTGAC TCACAGTACC	300
ACTGAAGTTA CTCAGCCAAG AACGAATACA CCAGTCAAAG AAGATTGGAA TGTCAGAATT	360
ACCAAGCTAC GGAAGCAAGT GGAAGAGATT TTTAATTTGA AATTTGCTCA AGCTCTTGGA	420
CTCACCGAGG CAGTAAAAGT ACCATATCCT GTGTTTGAAT CAAACCCGGA GTTCTTGTAT	480
GTGGAAGGCT TGCCAGAGGG GATTCCCTTC CGAAGCCCTA CCTGGTTTGG AATTCCACGA	540
CTTGAAAGGA TCGTCCACGG GAGTAATAAA ATCAAGTTCG TTGTTAAAAA ACCTGAACTA	600
GTTATTTTCT ACTTGCCCTC TGGGATGGCT AGTAAAATAA AACTAAAGC TTTGCAGTCC	660
CCCAAAAGAC CACGAAGTCC TGGGAGTAAT TCAAAGGTTT CTGAAATTGA GGTCACCGTG	720
GAAGGCCYTA ATAACAACAA TCCTCAAACC TCAGCTGTTC GAACCCCGAC CCAGACTAAC	780
GGTTCTAACG TTCCCTTCAA GCCACGAGGG AGAGAGTTTT CCTTTGGCCT GGAATGCCAA	840
AATCACGGAC CTAAAACAGA AAGTTGAAAA TCTCTTCAAT GAGAAATGTG GGGAAGCTCT	900
TGGCCTTAAA CAAGCTGTGA AGGTGCCGTT CGCGTTATTT GAGTC'TTCC CGGAAGACTT	960
TTATGTGGAA GGCTTACCTG AGGGTGTGCC ATTCCGAAGA CCATCGACTT TTGGCATTCC	1020
GAGGCTGGAG AAGATACTCA GAAACAAAGC CAAAATTAAG TTCATCATT AAGAGCCCGA	1080
AATGTTTGAG ACGGCGATTA AGGAGAGCAC CTCCTCTAAG AGCCCTCCCA GAAAAATAAA	1140
TTCATCACCC AATGTTAATA CTACTGCATC AGGTGTTGAA GACCTTAACA TCATTCAGGT	1200
GACAATTCCA GATGATGATA ATGAAAGACT CTCGAAAGTT GAAAAAGCTA GACAGCTAAG	1260
AGAACAAGTG AATGACCTCT TTAGTCGGAA ATTTGGTGAA GCTATTGGTA TGGGTTTTCC	1320

TGTGAAAGTT CCCTACAGGA AAATCACAAT TAACCCTGGC TGTGTGGTGG TTGATGGCAT 1380  
 GCCCCGGGG GTGTCCTTCA AAGCCCCCAG CTACCTGGAA ATCAGCTCCA TGAGAAGGAT 1440  
 CTTAGACTCT GCCGAGTTTA TCAAATTCAC GGTCATTAGA CCATTTCAG GACTTGTGAT 1500  
 TAATAACCAG CTGGTTGATC AGAGTGAGTC AAAAGGCCCC GTGATACAAG AATCAGCTGA 1560  
 ACCAAGCCAG TTGGAAGTTC CAGCCACAGA AGAAATAAAA GAGACTGATG GAAGCTCTCA 1620  
 GATCAAGCAA GAACCAGACC CCACGTGGTA GACCTCTTCC CTCCTAGGCT TAAAGTATCA 1680  
 GTGGTTGAGA AGAGCTTTTC GGACCTGTTA CTACCCCAAG CTGTGTAATA TACTTGTATA 1740  
 ACAGAAATAC CTTCTATACA AACCTTTTTT TCTACTTTTA GATAGAAATG TCTACTTTTT 1800  
 CAGCAGTTCT GTGAATTAAA GAGCAGAGTG ACTGTGGGTC TGAATGGCT GGTGTACTTG 1860  
 GGAATGTACT ATCAGGATTT TACAGCAATG CTGGGAAATG ACAGGGAAAA TGACAGGAAT 1920  
 GAATCTCACC AGATTTTTTA TGTACTCAGC AGAGCCTTGA GTTACGGTGT TTATTTTCCA 1980  
 ATCAAGTGAA GATATCTCCT ACTTCTCCTA CTGGAACATC TCAGCTTCTG CAGTGAAGAA 2040  
 AAATTCCTGT GATAGTTCAG TTCTTTAGTT TTTCTATTTG AAAAAAAAAA ATCATTTAAA 2100  
 TGATCCTTTG TTCACGGCTC TCCTTAATGA CTGAGTGAAC AGTTCCTATC TGTATATTTG 2160  
 ACTAAACCTT TTCCTAAGCT ATCTCTCATG GTTCCTATGT TTTTTTATCA TAATTAAAAG 2220  
 CAAAACCATT TGGATCACCT AACAGTCAGA GGTCAAGTATC TCAGCGTGTG AATTATAGAG 2280  
 GAAATACAGA GAGAACCTCT TCCACTTTTA CTTTTCGTCC AAATAAAATG CATGGTGTAC 2340  
 CAGAAGTTGA AGATCGGGTT GAGGATTGGG GCTAGCTCGA TGACACTAAG GCCCAACAT 2400  
 CGCGGGACCT GCTGTGGCGC GGATTCTTAG GAACGCTGTT CTAGCCGGCC CCCTCTCCAG 2460  
 GGGTCGCCGT GGCCGGCATT ATTTCTTAGT TCTTCTTGTA ACCCTGAGGT GCCAGCGCGG 2520  
 GGAGTGAGGA GGGGTCAGGG GGCTAAGGAT GCAACCTCTG ACGTTCTGCG CCTTCCTAGG 2580  
 AGAGTCTTAC ATGTGTTGAG ATTTACAAG CAATGCGAGT TGTAATAATC CAGCTCTACA 2640  
 AGAAGCTAGG CTCTGTGACG GCATAGTTTT CAGTAGCTTT ATCACAATAT TCACAATGGA 2700  
 GAATTATATG ACATGGTAGC AGAAATAGGC CCTTTTATGT GTTGCTTCTA TTTTACCTCA 2760  
 AATTGTAGAT ATAGGGTAAT CAATAAAATC CATCCATGCC TTTCAAAAAA AAAAAAAAAA 2820  
 AAAAAAAAAA AAAAA 2835

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 268 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Val Asp Gln Leu Phe Cys Lys Lys Phe Ala Glu Ala Leu Gly Ser
1           5           10           15

Thr Glu Ala Lys Ala Val Pro Tyr Gln Lys Phe Glu Ala His Pro Asn
          20           25           30

Asp Leu Tyr Val Glu Gly Leu Pro Glu Asn Ile Pro Phe Arg Ser Pro
          35           40           45

Ser Trp Tyr Gly Ile Pro Arg Leu Glu Lys Ile Ile Gln Val Gly Asn
          50           55           60

Arg Ile Lys Phe Val Ile Lys Arg Pro Glu Leu Leu Thr His Ser Thr
65           70           75           80

Thr Glu Val Thr Gln Pro Arg Thr Asn Thr Pro Val Lys Glu Asp Trp
          85           90           95

Asn Val Arg Ile Thr Lys Leu Arg Lys Gln Val Glu Glu Ile Phe Asn
          100          105          110

Leu Lys Phe Ala Gln Ala Leu Gly Leu Thr Glu Ala Val Lys Val Pro
          115          120          125

Tyr Pro Val Phe Glu Ser Asn Pro Glu Phe Leu Tyr Val Glu Gly Leu
          130          135          140

Pro Glu Gly Ile Pro Phe Arg Ser Pro Thr Trp Phe Gly Ile Pro Arg
          145          150          155          160

Leu Glu Arg Ile Val His Gly Ser Asn Lys Ile Lys Phe Val Val Lys
          165          170          175

Lys Pro Glu Leu Val Ile Ser Tyr Leu Pro Pro Gly Met Ala Ser Lys
          180          185          190

Ile Asn Thr Lys Ala Leu Gln Ser Pro Lys Arg Pro Arg Ser Pro Gly
          195          200          205

Ser Asn Ser Lys Val Pro Glu Ile Glu Val Thr Val Glu Gly Xaa Asn
          210          215          220

Asn Asn Asn Pro Gln Thr Ser Ala Val Arg Thr Pro Thr Gln Thr Asn
  
```

225                                      230                                      235                                      240  
 Gly Ser Asn Val Pro Phe Lys Pro Arg Gly Arg Glu Phe Ser Phe Gly  
    245                                      250                                      255  
 Leu Glu Cys Gln Asn His Gly Pro Lys Thr Glu Ser  
    260                                      265

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2705 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TAGGCCATGA AGGCCGCGCT TTTCGTCGAC TCTTACCGGT TGGCTGGGCC AGCTGCGCCG	60
CGGCTCACAG CTGACGATGG GGGACCCCAG CAAGCAGGAC ATCTTGACCA TCTTCAAGCG	120
CCTCCGCTCG GTGCCCCACTA ACAAGGTGTG TTTTGATTGT GGTGCCAAAA ATCCCAGCTG	180
GGCAAGCATA ACCTATGGAG TGTTCCTTTG CATTGATTGC TCAGGGTCCC ACCGGTCACT	240
TGGTGTTCAC TTGAGTTTTA TTCGATCTAC AGAGTTGGAT TCCAAGTGGT CATGGTTTCA	300
GTTGCGATGC ATGCAAGTCG GAGGAAACGC TAGTGCATCT TCCTTTTTTC ATCAACATGG	360
GTGTTCCACC AATGACACCA ATGCCAAGTA CAACAGTCGT GCTGCTCAGC TCTATAGGGA	420
GAAAATCAAA TCGCTCGCCT CTCAAGCAAC ACGGAAGCAT GGCACGTATC TGTGGCTTGA	480
TAGTTGTGTG GTTCCACCTT TGTCCCCTCC ACCAAAGGAG GAAGATTTTT TTGCCTCTCA	540
CGTTTCTCCT GAGGTGAGTG ACACAGCGTG GGCATCAGCA ATAGCAGAAC CATCTTCTTT	600
AACATCAAGG CCTGTGAAAA CCACTTTGGA AAATAATGAA GGTGGACAAG AGCAAGGACC	660
AAGTGTGGAA GGTCTTAATG TACCAACAAA GGCTACTTTA GAGGTATCCT CTATCATAAA	720
AAAGAAACCA AATCAAGCTA AAAAAGGCCT TGGGGCCAAA AAAGGAAGTT TGGGAGCTCA	780
GAAACTGGCA AACACATGCT TTAATGAAAT TGAAAAACAA GCTCAAGCTG CGGATAAAAT	840
GAAGGAGCAG AAAGACCTGG CCAAGGTGGT ATCTAAAGAA GAATCAATTG TTTCATCATT	900
ACGATTAGCC TATAAGGATC TTGAAATTCA AATGAAGAAA GACGAAAAGA TGAACATTAG	960

TGGCAAAAAA AATGTTGACT CAGACAGACT CGGCATGGGA TTTGGAAATT GCAGAAGTGT 1020  
 TATTTACAT TCAGTGACTT CAGATATGCA GACCATAGAG CAGGAATCAC CCATTATGGC 1080  
 AAAACCAAGA AAAAAGTATA ATGATGACAG TGACGATTCA TATTTTACTT CCAGCTCAAG 1140  
 TTACTTTGAC GAGCCAGTGG AGTTAAGGAG CAGTTCTTTC TCTAGCTGGG ATGACAGTTC 1200  
 AGATTCCAT TGGAAAAAG AGACCAGCAA AGATACTGAA ACAGTTCTGA AAACCACAGG 1260  
 CTATTCAGAC AGACCTACTG CTCGCCGCAA GCCAGATTAT GAGCCAGTTG AAAATACAGA 1320  
 TGAGGCCCAG AAGAAGTTTG GCAATGTCAA GGCCATTTCA TCAGATATGT ATTTTGGAAG 1380  
 ACAATCCCAG GCTGATTATG AGACCAGGGC CCGCCTAGAG AGGCTGTCGG CAAGTTCCTC 1440  
 CATAAGCTCG GCTGATCTGT TCGAGGAGCC GAGGAAGCAG CCAGCAGGGA ACTACAGCCT 1500  
 GTCCAGTGTG CTGCCCCAAG CCCCCGACAT GGCGCAGTTC AAGCAGGGAG TGAGATCGGT 1560  
 TGCTGGAAAA CTCTCCGTCT TTGCTAATGG AGTCGTGACT TCAATTCAGG ATCGCTACGG 1620  
 TTCTTAATAC TGAAGTCATG ATGTGTATTT CCTGGAGAAA TTCCTCTTTA AATGAACAAG 1680  
 TAACCACATC TCAGGCGGCA GTGAAGTCCA GATAGTTTTC CAGATTGTTT TGCTACTTTT 1740  
 TCATATGGTA TATGTTTCTG ATTTTAAATA TTTCTTTTGA GAAATTCTGA GTTCTGATGT 1800  
 AGGAGCTTTC CTGTGATTTT TGTTCACGT TCCTTCCTGT CACACCCTCC TTGGCGTCT 1860  
 CTGTGTATAT CCTTGCTTTA TTTTCTTGGA ACCTTTGATT TCAACACTGA GGGCCTGGAG 1920  
 ACCTCGGCTC CTCCTGCTCC TGAACCAGGA GGCTTCATGT GGGGGAGGAG GAGAGGTCTC 1980  
 CATGTGACAC ATGGGCTCAG GGCTGCCAGA ATCAGCGGAT GCTGGATGGG CCTGCAGAAA 2040  
 CAACACTCAC CACACACACT TCCTTCAAAA GACCAAAAGT GACTGGTGTC TCGTGTGACA 2100  
 GATTGCTTCA TTTATGTTTC TACATAGTAA GGTGACTGCC AAATAATATT TGAAGTCATC 2160  
 TGTCTCTTTG TAAATTATTT TATATGACCT ATAAATTTAA AAATGTTTTT CAGTGAGTGC 2220  
 TTTTAACAAA CTTAAGCTTC TGCCCTGCCA AGGGAATTAA TGTTATCTTG TGAAAGGTGT 2280  
 TGCTGTTTGA ATTGATGAGA AATGGAAGAT GAGAACTCCC TAAGAGTTCT CATAATAAAT 2340  
 CATCTCATCA CAAATCAATA CGGTATACAG AGTTAAAGTG GAATGAGGTA AGAAGATACA 2400  
 GCTACAGAAA ATAGTTGCGT GTATGGGAGA ACAGTCATTG TAATTGGGTA GTTTTGTAA 2460  
 TAAATATTTT TAAATCTTGC TTTTCAGAAA TTACCGAATG TGTATAAACA AATAAGAAA 2520  
 AATAATTTAG CTGTGTTTTA GACAGCATTA GAATATATTG TTCAGCACAG TAAAATATAT 2580  
 TTGAAATTTG ATAAGCCAAA AATGTGGTTT TGAATGAATA TTTTGTGAAT CTTTCTTAAA 2640

AGCTCAAATT TGTAAGCTTC TAAATAGAAT AAACACTTGC AGCAGATGGA AAAAAAAAAA 2700

AAAAA 2705

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 516 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Gly	Asp	Pro	Ser	Lys	Gln	Asp	Ile	Leu	Thr	Ile	Phe	Lys	Arg	Leu	1	5	10	15
Arg	Ser	Val	Pro	Thr	Asn	Lys	Val	Cys	Phe	Asp	Cys	Gly	Ala	Lys	Asn	20	25	30	
Pro	Ser	Trp	Ala	Ser	Ile	Thr	Tyr	Gly	Val	Phe	Leu	Cys	Ile	Asp	Cys	35	40	45	
Ser	Gly	Ser	His	Arg	Ser	Leu	Gly	Val	His	Leu	Ser	Phe	Ile	Arg	Ser	50	55	60	
Thr	Glu	Leu	Asp	Ser	Asn	Trp	Ser	Trp	Phe	Gln	Leu	Arg	Cys	Met	Gln	65	70	75	80
Val	Gly	Gly	Asn	Ala	Ser	Ala	Ser	Ser	Phe	Phe	His	Gln	His	Gly	Cys	85	90	95	
Ser	Thr	Asn	Asp	Thr	Asn	Ala	Lys	Tyr	Asn	Ser	Arg	Ala	Ala	Gln	Leu	100	105	110	
Tyr	Arg	Glu	Lys	Ile	Lys	Ser	Leu	Ala	Ser	Gln	Ala	Thr	Arg	Lys	His	115	120	125	
Gly	Thr	Asp	Leu	Trp	Leu	Asp	Ser	Cys	Val	Val	Pro	Pro	Leu	Ser	Pro	130	135	140	
Pro	Pro	Lys	Glu	Glu	Asp	Phe	Phe	Ala	Ser	His	Val	Ser	Pro	Glu	Val	145	150	155	160
Ser	Asp	Thr	Ala	Trp	Ala	Ser	Ala	Ile	Ala	Glu	Pro	Ser	Ser	Leu	Thr	165	170	175	
Ser	Arg	Pro	Val	Glu	Thr	Thr	Leu	Glu	Asn	Asn	Glu	Gly	Gly	Gln	Glu	180	185	190	

Gln Gly Pro Ser Val Glu Gly Leu Asn Val Pro Thr Lys Ala Thr Leu  
 195 200 205  
 Glu Val Ser Ser Ile Ile Lys Lys Lys Pro Asn Gln Ala Lys Lys Gly  
 210 215 220  
 Leu Gly Ala Lys Lys Gly Ser Leu Gly Ala Gln Lys Leu Ala Asn Thr  
 225 230 235 240  
 Cys Phe Asn Glu Ile Glu Lys Gln Ala Gln Ala Ala Asp Lys Met Lys  
 245 250 255  
 Glu Gln Lys Asp Leu Ala Lys Val Val Ser Lys Glu Glu Ser Ile Val  
 260 265 270  
 Ser Ser Leu Arg Leu Ala Tyr Lys Asp Leu Glu Ile Gln Met Lys Lys  
 275 280 285  
 Asp Glu Lys Met Asn Ile Ser Gly Lys Lys Asn Val Asp Ser Asp Arg  
 290 295 300  
 Leu Gly Met Gly Phe Gly Asn Cys Arg Ser Val Ile Ser His Ser Val  
 305 310 315 320  
 Thr Ser Asp Met Gln Thr Ile Glu Gln Glu Ser Pro Ile Met Ala Lys  
 325 330 335  
 Pro Arg Lys Lys Tyr Asn Asp Asp Ser Asp Asp Ser Tyr Phe Thr Ser  
 340 345 350  
 Ser Ser Ser Tyr Phe Asp Glu Pro Val Glu Leu Arg Ser Ser Ser Phe  
 355 360 365  
 Ser Ser Trp Asp Asp Ser Ser Asp Ser Tyr Trp Lys Lys Glu Thr Ser  
 370 375 380  
 Lys Asp Thr Glu Thr Val Leu Lys Thr Thr Gly Tyr Ser Asp Arg Pro  
 385 390 395 400  
 Thr Ala Arg Arg Lys Pro Asp Tyr Glu Pro Val Glu Asn Thr Asp Glu  
 405 410 415  
 Ala Gln Lys Lys Phe Gly Asn Val Lys Ala Ile Ser Ser Asp Met Tyr  
 420 425 430  
 Phe Gly Arg Gln Ser Gln Ala Asp Tyr Glu Thr Arg Ala Arg Leu Glu  
 435 440 445  
 Arg Leu Ser Ala Ser Ser Ser Ile Ser Ser Ala Asp Leu Phe Glu Glu  
 450 455 460  
 Pro Arg Lys Gln Pro Ala Gly Asn Tyr Ser Leu Ser Ser Val Leu Pro  
 465 470 475 480  
 Asn Ala Pro Asp Met Ala Gln Phe Lys Gln Gly Val Arg Ser Val Ala



485

490

495

Gly Lys Leu Ser Val Phe Ala Asn Gly Val Val Thr Ser Ile Gln Asp  
500 505 510

Arg Tyr Gly Ser  
515

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1414 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTTCCAGCC CGAGGGCATC GCGCTGGCCT ACGGCAGCCT CCTGCTCATG GCGCTGCTGC	60
CCATCTTCTT CGGCGCCCTG CGCTCCGTAC GCTGCGCCCG CGGCAAGAAT GCTTCAGACA	120
TGCCTGAAAC AATCACCAGC CGGGATGCCG CCCGCTTCCC CATCATCGCC AGCTGCACAC	180
TCTTGGGGCT CTACCTCTTT TTCAAAATAT TCTCCCAGGA GTACATCAAC CTCCTGCTGT	240
CCATGTATTT CTTGCTGCTG GGAATCCTGG CCCTGTCCCA CACCATCAGC CCCTTCATGA	300
ATAAGTTTTT TCCAGCCAGC TTTCCAAATC GACAGTACCA GCTGCTCTTC ACACAGGGTT	360
CTGGGGAAAA CAAGGAAGAG ATCATCAATT ATGAATTTGA CACCAAGGAC CTGGTGTGCC	420
TGGGCTGAG CAGCATCGTT GGCGTCTGGT ACCTGCTGAG GAAGCACTGG ATTGCCAACA	480
ACCTTTTGG CCTGGCCTTC TCCCTTAATG GAGTAGAGCT CCTGCACCTC AACAATGTCA	540
GCACTGGCTG CATCCTGCTG GGCGGACTCT TCATCTACGA TGTCTTCTGG GTATTTGGCA	600
CCAATGTGAT GGTGACAGTG GCCAAGTCCT TCGAGGCACC AATAAAATTG GTGTTTCCCC	660
AGGATCTGCT GGAGAAAGGC CTCGAAGCAA ACAACTTTGC CATGCTGGGA CTTGGAGATG	720
TCGTCATTCC AGGGATCTTC ATTGCCTTGC TGCTGCGCTT TGACATCAGC TTGAAGAAGA	780
ATACCCACAC CTACTTCTAC ACCAGCTTTG CAGCCTACAT CTTGGGCCTG GGCCTTACCA	840
TCTTCATCAT GCACATCTTC AAGCATGCTC AGCCTGCCCT CCTATACCTG GTCCCCGCCT	900
GCATCGGTTT TCCTGTCCTG GTGGCGCTGG CCAAGGGAGA AGTGACAGAG ATGTTCA GTT	960

ATGAGGAGTC AAATCCTAAG GATCCAGCGG CAGTGACAGA ATCCAAAGAG GGAACAGAGG 1020  
 CATCAGCATC GAAGGGGCTG GAGAAGAAAG AGAAATGATG CAGCTGGTGC CCGAGCCTCT 1080  
 CAGGGCCAGA CCAGACAGAT GGGGGCTGGG CCCACACAGG CGTGCACCGG TAGAGGGCAC 1140  
 AGGAGGCCAA GGGCAGCTCC AGGACAGGGC AGGGGGCAGC AGGATACCTC CAGCCAGGCC 1200  
 TCTGTGGCCT CTGTTTCCTT CTCCCTTTCT TGGCCCTCCT CTGCTCCTCC CCACACCCTG 1260  
 CAGGCAAAAG AAACCCCCAG CTTCCCCCT CCCC GGAGC CAGGTGGGAA AAGTGGGTGT 1320  
 GATTTTTAGA TTTTGTATTG TGGACTGATT TTGCCTCACA TTAAAACTC ATCCCATGGC 1380  
 CAGGGCGGGC CACTGTAAAA AAAAAAAAAA AAAA 1414

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 336 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met	Ala	Leu	Leu	Pro	Ile	Phe	Phe	Gly	Ala	Leu	Arg	Ser	Val	Arg	Cys
1				5					10					15	
Ala	Arg	Gly	Lys	Asn	Ala	Ser	Asp	Met	Pro	Glu	Thr	Ile	Thr	Ser	Arg
			20					25						30	
Asp	Ala	Ala	Arg	Phe	Pro	Ile	Ile	Ala	Ser	Cys	Thr	Leu	Leu	Gly	Leu
			35				40							45	
Tyr	Leu	Phe	Phe	Lys	Ile	Phe	Ser	Gln	Glu	Tyr	Ile	Asn	Leu	Leu	Leu
			50				55					60			
Ser	Met	Tyr	Phe	Phe	Val	Leu	Gly	Ile	Leu	Ala	Leu	Ser	His	Thr	Ile
			65			70				75				80	
Ser	Pro	Phe	Met	Asn	Lys	Phe	Phe	Pro	Ala	Ser	Phe	Pro	Asn	Arg	Gln
				85					90					95	
Tyr	Gln	Leu	Leu	Phe	Thr	Gln	Gly	Ser	Gly	Glu	Asn	Lys	Glu	Glu	Ile
			100					105					110		
Ile	Asn	Tyr	Glu	Phe	Asp	Thr	Lys	Asp	Leu	Val	Cys	Leu	Gly	Leu	Ser
			115				120						125		

Ser Ile Val Gly Val Trp Tyr Leu Leu Arg Lys His Trp Ile Ala Asn  
 130 135 140  
 Asn Leu Phe Gly Leu Ala Phe Ser Leu Asn Gly Val Glu Leu Leu His  
 145 150 155 160  
 Leu Asn Asn Val Ser Thr Gly Cys Ile Leu Leu Gly Gly Leu Phe Ile  
 165 170 175  
 Tyr Asp Val Phe Trp Val Phe Gly Thr Asn Val Met Val Thr Val Ala  
 180 185 190  
 Lys Ser Phe Glu Ala Pro Ile Lys Leu Val Phe Pro Gln Asp Leu Leu  
 195 200 205  
 Glu Lys Gly Leu Glu Ala Asn Asn Phe Ala Met Leu Gly Leu Gly Asp  
 210 215 220  
 Val Val Ile Pro Gly Ile Phe Ile Ala Leu Leu Leu Arg Phe Asp Ile  
 225 230 235 240  
 Ser Leu Lys Lys Asn Thr His Thr Tyr Phe Tyr Thr Ser Phe Ala Ala  
 245 250 255  
 Tyr Ile Phe Gly Leu Gly Leu Thr Ile Phe Ile Met His Ile Phe Lys  
 260 265 270  
 His Ala Gln Pro Ala Leu Leu Tyr Leu Val Pro Ala Cys Ile Gly Phe  
 275 280 285  
 Pro Val Leu Val Ala Leu Ala Lys Gly Glu Val Thr Glu Met Phe Ser  
 290 295 300  
 Tyr Glu Glu Ser Asn Pro Lys Asp Pro Ala Ala Val Thr Glu Ser Lys  
 305 310 315 320  
 Glu Gly Thr Glu Ala Ser Ala Ser Lys Gly Leu Glu Lys Lys Glu Lys  
 325 330 335

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1583 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGACGAGCCT TTCTTATTTC TTTACTCAAC CTCTTTGATG ACACAGCAAA AACAGACGTG 60  
 ACTATGCTCT TGTATATAGC AGACAATCTA GCCTGTTTTTC CATACCAGAC ACAGGAAGAG 120  
 CCGTTGTTTA TAATGCATCA TATAGACATT ACACTCTCAG TTTCTGGTAG TAACCTACTG 180  
 CAGTCATTCA AGGAGTCTAT GGTAAGGAC AAAAGGAAAG AGAGAAAATC ATCACCTAGT 240  
 AAGGAAAATG AGTCAAGCGA CAGTGAAGAA GAAGTTTCCA GGCCTCGGAA GTCACGGAAA 300  
 CGTGTAGATT CAGATTCAGA TTCAGATTCA GAAGACGATA TAAATTCAGT GATGAAATGT 360  
 TTGCCAGAAA ATTCAGCTCC TTTAATCGAA TTTGCAAATG TGTCCCAGGG TATTTTATTA 420  
 CTTCTCATGT TAAAACAACA TTTGAAGAAT CTTTGTGGAT TTTCTGATAG TAAATTCAG 480  
 AAGTACTCTC CATCTGAATC TGCAAAAGTA TATGATAAAG CGATAAACCG AAAACAGGA 540  
 GTTCATTTTC ATCCAAAACA AACACTGGAC TTCTTGCGGA GTGACATGGC TAATTCCTCAA 600  
 ATCACAGAAG AGGTGAAAAG GAGTATAGTA AAACAGTATC TAGATTTCAA ACTTCTCATG 660  
 GAACATCTGG ACCCTGATGA AGAAGAAGAA GAAGGGGAGG TTTCAGCTAG CACAAATGCT 720  
 CGGAACAAAG CAATTACCTC ACTGCTTGA GGAGGCAGCC CTAAAAATAA TACAGCAGCA 780  
 GAGACAGAAG ATGATGAAAG TGATGGGGAG GATAGAGGAG GAGGCACTTC AGGGTCATTG 840  
 AGAAGGTCAA AACGAAATTC AGACTCTACG GAGTTGGCAG CACAGATGAA TGAAAGTGTT 900  
 GACGTCATGG ATGTCATCGC TATTTGCTGT CCAAAGTACA AAGATCGACC ACAAATTGCA 960  
 AGAGTAGTGC AGAAAACCAG CAGTGGCTTC AGTGTTTCAGT GGATGGCAGG CTCCTACAGT 1020  
 GGCTCCTGGA CTGAGGCTAA GCGCCGTGAT GGCCGCAAAC TGGTGCCTTG GGTAGACACT 1080  
 ATTAAAGAGT CAGACATTAT TTACAAAAAA ATTGCTCTAA CGAGTGCTAA TAAGCTGACT 1140  
 AATAAGTTG TTCAGACTTT ACGATCCCTG TATGCCGCCA AGGATGGGAC TTCCAGCTAA 1200  
 TGAATTTGTA CATGCAGCCA AATTACAGG AATTTTTTTA AAAGGCAGAA AAACCTGAAA 1260  
 TACCAACATT CTGGCAAAAA AAAATCAGTT TTATGAAGAG TAAGTGGAAC CTGGGATGCA 1320  
 GGAACAAAAG AAGGAAATGT TGGGCAAACA TTTTGTGGG AGCTCCCTTC GCTGTTGTGC 1380  
 AGCAGAAACA GATTCTCAGT TCATTTTAC TCCCACTGTA TTATAGTTTA AAAAAATTG 1440  
 TTTATATCTT GGAAAAAAA ACTTTCTGTT TAAAAAAAAT AAACAAGTGA ATGTTGGAAA 1500  
 TTAGTCTGTT AATGTTCTTA ATAAAGTGTT CTTGGAGTTT TAAAAAAA AAATAAAAAA 1560  
 AAAAAAAA AAAAAAAA AAA 1583

(2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 378 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Leu Leu Tyr Ile Ala Asp Asn Leu Ala Cys Phe Pro Tyr Gln Thr
1           5           10           15

Gln Glu Glu Pro Leu Phe Ile Met His His Ile Asp Ile Thr Leu Ser
          20           25           30

Val Ser Gly Ser Asn Leu Leu Gln Ser Phe Lys Glu Ser Met Val Lys
          35           40           45

Asp Lys Arg Lys Glu Arg Lys Ser Ser Pro Ser Lys Glu Asn Glu Ser
          50           55           60

Ser Asp Ser Glu Glu Glu Val Ser Arg Pro Arg Lys Ser Arg Lys Arg
65           70           75           80

Val Asp Ser Asp Ser Asp Ser Asp Ser Glu Asp Asp Ile Asn Ser Val
          85           90           95

Met Lys Cys Leu Pro Glu Asn Ser Ala Pro Leu Ile Glu Phe Ala Asn
          100          105          110

Val Ser Gln Gly Ile Leu Leu Leu Leu Met Leu Lys Gln His Leu Lys
          115          120          125

Asn Leu Cys Gly Phe Ser Asp Ser Lys Ile Gln Lys Tyr Ser Pro Ser
          130          135          140

Glu Ser Ala Lys Val Tyr Asp Lys Ala Ile Asn Arg Lys Thr Gly Val
          145          150          155          160

His Phe His Pro Lys Gln Thr Leu Asp Phe Leu Arg Ser Asp Met Ala
          165          170          175

Asn Ser Lys Ile Thr Glu Glu Val Lys Arg Ser Ile Val Lys Gln Tyr
          180          185          190

Leu Asp Phe Lys Leu Leu Met Glu His Leu Asp Pro Asp Glu Glu Glu
          195          200          205

Glu Glu Gly Glu Val Ser Ala Ser Thr Asn Ala Arg Asn Lys Ala Ile
          210          215          220

```

Thr Ser Leu Leu Gly Gly Gly Ser Pro Lys Asn Asn Thr Ala Ala Glu  
 225 230 235 240  
 Thr Glu Asp Asp Glu Ser Asp Gly Glu Asp Arg Gly Gly Gly Thr Ser  
 245 250 255  
 Gly Ser Leu Arg Arg Ser Lys Arg Asn Ser Asp Ser Thr Glu Leu Ala  
 260 265 270  
 Ala Gln Met Asn Glu Ser Val Asp Val Met Asp Val Ile Ala Ile Cys  
 275 280 285  
 Cys Pro Lys Tyr Lys Asp Arg Pro Gln Ile Ala Arg Val Val Gln Lys  
 290 295 300  
 Thr Ser Ser Gly Phe Ser Val Gln Trp Met Ala Gly Ser Tyr Ser Gly  
 305 310 315 320  
 Ser Trp Thr Glu Ala Lys Arg Arg Asp Gly Arg Lys Leu Val Pro Trp  
 325 330 335  
 Val Asp Thr Ile Lys Glu Ser Asp Ile Ile Tyr Lys Lys Ile Ala Leu  
 340 345 350  
 Thr Ser Ala Asn Lys Leu Thr Asn Lys Val Val Gln Thr Leu Arg Ser  
 355 360 365  
 Leu Tyr Ala Ala Lys Asp Gly Thr Ser Ser  
 370 375

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2353 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TCCCGTCGCG CTCAGCATCC TCAATCATCC GCAGGCTGAT GCGGTCCTTC ACGCTCCCGC 60  
 CCGCGTTGAA GAACTCACAC TTGGCCAAGA GCTCACACTT CAGGCCGAAC TTCTTCCCAA 120  
 TCTTGTTGAT TCTGACCATA GGGGTGTCCC CGATTTTCTT CAGAATATCT GGCAAGATTT 180  
 TTGGAGATTT TGCCGGGGCA GTGTGGTGAT GTGGGGACTC GGAGGCAGGC CGGCCAGCT 240  
 GCCAGGTGCA CCTGCTCGGA GCATCGGGCC GGATCCACAG GGGCTCCTTG GCTTCCTTAT 300

CCTCTGGGGA CCCCTTCTCC AGGCTCCCCCT TCGCCGAGTG TGGCCCTGAG CGGTGGGGGC	360
AGCCTGTGGG CCCCACTTCT GCCTGGGGGG TCTCAGAAGG CATGCTGGGA CCGGACAGAT	420
GGTGCTGACA GTTCAGCACC GCTCAAAC TC TGTAAGGAGA CGTGACAACA GAATCTGTCA	480
TGGGATCCTG GAATAGAAAA GGGCAGTGAG GAGAAAAC TC TGCAAATCTG ACTCAAGCGC	540
GGACTTTGGT TAGGCCTCTT TGGACAGTAC TTTGAGAATT ATGGATCTCC GGATATTAGT	600
AGCAAGGAAG TTTGTAGGAG CAGCAAATTA TCGGGAGAAG ATTCATAGTA CTTTGACTCC	660
ATGTGGGACT TTTCTGTTTG CTGGAAGTGA GGATGGTATA GTGTATGTTT GGAACCCAGA	720
AACAGGAGAA CAAGTAGCCA TGTATTCTGA CTTGCCATTC AAGTCACCCA TTCGAGACAT	780
TTCTTATCAT CCATTTGAAA ATATGGTTGC ATTCTGTGCA TTTGGGCAAA ATGAGCCAAT	840
TCTTCTGTAT ATTTACGATT TCCATGTTGC CCAGCAGGAG GCTGAAATGT TCAAACGCTA	900
CAATGGAACA TTTCCATTAC CTGGAATACA CCAAAGTCAA GATGCCCTAT GTACCTGTCC	960
AAACTACCC CATCAAGGCT CTTTTCAGAT TGATGAATTT GTCCACACTG AAAGTTCTTC	1020
AACGAAGATG CAGCTAGTAA AACAGAGGCT TGAAACTGTC ACAGAGGTGA TACGTTCTTG	1080
TGCTGCAAAA GTCAACAAAA ATCTCTCATT TACTTCACCA CCAGCAGTTT CCTCACAACA	1140
GTCTAAGTTA AAGCAGTCAA ACATGCTGAC CGCTCAAGAG ATTCTACATC AGTTTGGTTT	1200
CACTCAGACC GGGATTATCA GCATAGAAAG AAAGCCTTGT AACCATCAGG TAGATACAGC	1260
ACCAACGGTA GTGGCTCTTT ATGACTACAC AGCGAATCGA TCAGATGAAC TAACCATCCA	1320
TCGCGGAGAC ATTATCCGAG TGTTTTTCAA AGATAATGAA GACTGGTGGT ATGGCAGCAT	1380
AGGAAAGGGA CAGGAAGGTT ATTTTCCAGC TAATCATGTG GCTAGTGAAA CACTGTATCA	1440
AGAACTGCCT CCTGAGATAA AGGAGCGATC CCCTCCTTTA AGCCCTGAGG AAAAACTAA	1500
AATAGAAAAA TCTCCAGCTC CTCAAAGGT AAAATAAAAC AAAACAGCTC ACAGAGACAC	1560
CTCCTTCTTC CCCCTGCTTC TGCCTCCATG AGTAACTACC TATATGACGT GCTGCTGCTG	1620
AGACCAAGGA ATGAGTGAGT AAAGGTGTTT GGAAGTCAAA TATTGGTCCT AGTTAAGTAA	1680
GTGTCCTTTT CAAGTAAGTG GTAAACTTTG TAATGTGGAC CCCCTTCTAA CTTAGGATTC	1740
ATATAATTTG AACACAGGCT AAGCTGCCTC TGCATCATTA AAAAGGATTT GGGTATGCCA	1800
GTGATAGGAA TACACAGTAG AAAAGATGGC AGGGACCATG TTTAACATAC CTAAGACAAA	1860
CTTTCTGAGG ATATCACAGC ATCCATTTCA CATGAACAGA CTTAGCAGAA TTGGCAAGAC	1920
CTGCACATGG CAACATATTA TTCATTTTTC AGGATAACAA AAAATAATTA AAAGTAACAT	1980

TGGCTTAACA CTGGTATATA GTTCAGTATG ACTTCTCTCC CCTCAACCTC TGGAATAGGT 2040  
TCAAATTGAT GAGAATTTCT GATTAGAGCC CTTAATGTTG AGTTTTTTGA AAAGTTTAT 2100  
CAAGTTTCAT ATATACCTAT ATTGATGGTA AGTTGCTGGT CTTGCCATGG GCAAAGAGAG 2160  
AAAAATGATA YTGAGACCTT GTAAAGAATA GYTGGACATG GAGGCGCACA CCTGTAATCC 2220  
CAGCAAYTCA GAAGGCTGAG GCGGGAGGAT TGCTTGAGCC CAGGAGTTCA AGGCTGCAGT 2280  
GAGCTACGAT CATGCCACTG TACTCCAGCC TGAGCAACAA AGCATGACCC AATCTCTTAA 2340  
AAAAAAAAAA AAA 2353

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 318 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Asp Leu Arg Ile Leu Val Ala Arg Lys Phe Val Gly Ala Ala Asn  
1 5 10 15  
Tyr Arg Glu Lys Ile His Ser Thr Leu Thr Pro Cys Gly Thr Phe Leu  
20 25 30  
Phe Ala Gly Ser Glu Asp Gly Ile Val Tyr Val Trp Asn Pro Glu Thr  
35 40 45  
Gly Glu Gln Val Ala Met Tyr Ser Asp Leu Pro Phe Lys Ser Pro Ile  
50 55 60  
Arg Asp Ile Ser Tyr His Pro Phe Glu Asn Met Val Ala Phe Cys Ala  
65 70 75 80  
Phe Gly Gln Asn Glu Pro Ile Leu Leu Tyr Ile Tyr Asp Phe His Val  
85 90 95  
Ala Gln Gln Glu Ala Glu Met Phe Lys Arg Tyr Asn Gly Thr Phe Pro  
100 105 110  
Leu Pro Gly Ile His Gln Ser Gln Asp Ala Leu Cys Thr Cys Pro Lys  
115 120 125  
Leu Pro His Gln Gly Ser Phe Gln Ile Asp Glu Phe Val His Thr Glu  
130 135 140



Ser Ser Ser Thr Lys Met Gln Leu Val Lys Gln Arg Leu Glu Thr Val  
 145 150 155 160  
 Thr Glu Val Ile Arg Ser Cys Ala Ala Lys Val Asn Lys Asn Leu Ser  
 165 170 175  
 Phe Thr Ser Pro Pro Ala Val Ser Ser Gln Gln Ser Lys Leu Lys Gln  
 180 185 190  
 Ser Asn Met Leu Thr Ala Gln Glu Ile Leu His Gln Phe Gly Phe Thr  
 195 200 205  
 Gln Thr Gly Ile Ile Ser Ile Glu Arg Lys Pro Cys Asn His Gln Val  
 210 215 220  
 Asp Thr Ala Pro Thr Val Val Ala Leu Tyr Asp Tyr Thr Ala Asn Arg  
 225 230 235 240  
 Ser Asp Glu Leu Thr Ile His Arg Gly Asp Ile Ile Arg Val Phe Phe  
 245 250 255  
 Lys Asp Asn Glu Asp Trp Trp Tyr Gly Ser Ile Gly Lys Gly Gln Glu  
 260 265 270  
 Gly Tyr Phe Pro Ala Asn His Val Ala Ser Glu Thr Leu Tyr Gln Glu  
 275 280 285  
 Leu Pro Pro Glu Ile Lys Glu Arg Ser Pro Pro Leu Ser Pro Glu Glu  
 290 295 300  
 Lys Thr Lys Ile Glu Lys Ser Pro Ala Pro Gln Lys Val Lys  
 305 310 315

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2567 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGCCACCTA CACACCCAGC TCCCACCCAT GGAGAGCACC CCAGTCCTGT TCCTCCTGCC 60  
 CTGGACCCCTG GCCAACTCTG CCACAAGCTC TACCCTCGGT ACAACAGGCT CTGTCCCCAC 120  
 ATCTACAGAC CCTGCCCCAT CTGCACACCT AGACTCAGTT CATAAGTCCA CAGACTCTGG 180  
 CCCTTCAGAA CTGCCAGGCC CCACTCACAC CACTACAGGC TCTACCTATA GTGCCATTAC 240

CACTACCCAC AGTGCTCCAA GCCCCCTCAC TCACACTACT ACAGGCTCCA CCCACAAGCC 300  
 CATAATCTCT ACCCTTACTA CTACAGGCCC TACCCTCAAT ATCATAGGCC CAGTCCAGAC 360  
 TACCACAAGC CCCACCCACA CTATGCCAAG CCCTACCCAT ACCACAGCAA GCCCCACTCA 420  
 TACTTCCACA AGCCCCACCC ATACCCCCAC AAGTCCCACC CACAAAACCA GTATGTCACC 480  
 TCCCACCACT ACAAGTCCTA CCCCCAGTGG TATGGGCTTA GTCCAGACTG CCACAAGTCC 540  
 CACCCATCCT ACCACAAGCC CCACCCATCC CACCACAAGC CCCATCCTTA TAAATGTAAG 600  
 CCCTTCCACT TCTCTAGAAC TTGCTACCCT CTCCAGCCCC TCCAAACACT CAGACCCAC 660  
 CCTCCCAGGC ACTGACTCCC TTCCCTGTAG TCCCCCAGTC TCCAATTCCT AACTCAGGC 720  
 AGACCCATATG GCCCCAGAA CTCCCCACCC AAGTCCTGCC CATTCCAGTA GGAAACCCCT 780  
 CACAAGCCCT GCCCCAGATC CCTCAGAGTC TACGGTTCAG AGCTAAGCC CCACTCCCCTC 840  
 ACCCCCAACC CCTGCACCCC AGCATTCAGA CCTTTGCCTG GCCATGGCTG TCCAGACCCC 900  
 AGTCCCAACG GCAGCCGGAG GGTCTGGGA CAGGAGCCTG GAGGAGGCAC TGGGGGCCCT 960  
 AATGGCTGCC CTGGATGACT ACCGTGGCCA GTTTCCTGAG CTGCAGGGCC TGGAGCAGGA 1020  
 GGTGACCCGC CTAGAAAGTC TGCTCATGCA GAGACAAGGT CTGACTCGCA GCCGGCCTC 1080  
 CAGTCTCAGC ATCACTGTGG AGCATGCCTT GGAGAGCTTC AGCTTCCTCA ATGAAAACGA 1140  
 AGATGAAGAC AATGATGTTT CTGGGGACAG GCCTCCAAGC AGCCCGGAGG CTGGGGCTGA 1200  
 GGACAGCATA GACTCACCCA GTGCCCCGCC CCTCAGCACG GGGTGTCCAG CTCTGGATGC 1260  
 TGCCTTGGTC CGGCACCTGT ACCACTGCAG TCGCCTCCTG CTGAAACTGG GCACATTTGG 1320  
 GCCCCTGCGC TGCCAGGAGG CATGGGCCCT GGAGCGGCTG CTGCGGGAAG CCCGAGTACT 1380  
 GGAGGCAGTA TGCGAGTTCA GCAGGCGGTG GGAGATCCCG GCCAGCTCTG CCCAGGAAGT 1440  
 GGTGCAGTTC TCGGCCTCTC GGCCTGGCTT CCTGACCTTC TGGGACCAGT GCACAGAGAG 1500  
 ACTCAGCTGC TTCCTCTGCC CGGTGGAGCG GGTGCTTCTC ACCTTCTGCA ACCAGTATGG 1560  
 TGCCCGCCTC TCCCTGCGCC AGCCAGGCTT GGCTGAGGCT GTGTGTGTGA AGTTCCTGGA 1620  
 GGATGCCCTG GGGCAGAAGC TGCCCAGAAG GCCCCAGCCA GGGCCTGGAG AGCAGCTCAC 1680  
 AGTCTTCCAG TTCTGGAGTT TTGTGGAAAC CTTGGACAGC CCCACCATGG AGGCCTACGT 1740  
 GACTGAGACC GCTGAGGAGG TGCTACTGGT GCGGAATCTG AACTCGGATG ATCAGGCTGT 1800  
 TGTGCTGAAG GCCCTGAGAT TGGCGCCCGA GGGGCGTCTG CGAAGGGACG GGCTGCGGGC 1860  
 CCTCAGCTCC CTGCTCGTCC ATGGCAACAA CAAGGTCATG GCTGCTGTCA GCACCCAGCT 1920

CCGGAGCCTG TCACTGGGCC CTACCTTCCG GGAGAGGGCC CTCCTGTGCT TCCTGGACCA 1980  
 GCTGGAGGAT GAGGACGTGC AGACTCGAGT GGCTGGCTGC CTGGCCCTAG GCTGCATCAA 2040  
 GGCTCCCGAG GGCATTGAGC CCCTGGTGTA CCTCTGCCAA ACTGACACAG AAGCTGTGAG 2100  
 GGAAGCTGCC CGGCAAAGCC TACAGCAGTG TGGAGAAGAG GGACAGTCTG CCCATCGACG 2160  
 GCTGGAGGAG TCCCTGGACG CCCTGCCCCG CATCTTTGGG CCTGGCAGCA TGGCCAGCAC 2220  
 AGCATTCTAA ACTATTCACC CATGGGTTCC TGGTGCCCTT TTCCCCCAC TTTCAGGGCT 2280  
 CACCAGGCAC TGGCAGGGAG GGTAAGGGCT GGCTCCAGAT ACCCCTCCCC CACAGATTCC 2340  
 TAGCAATGAA AATCTAATAT ATTCTTCTGT TGCCCTGGG GTTGGAGAGT CAGTGCCTGC 2400  
 AGTCAAGTGC CTCCCAGCCT CGGCTCAGCA CATCCCTTGC CACAAATCAG TGTCTGGGGC 2460  
 TTGGCCACCC TGCCGCTGCC CAGCCACATC CCTTGGTTTT GTATTTTATT TACAGAGTTT 2520  
 TACAGAAAAT AAAAAAGCAA AATGTCTTTC CTAAAAAAA AAAAAA 2567

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 733 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Glu Ser Thr Pro Val Leu Phe Leu Leu Pro Trp Thr Leu Ala Asn  
 1 5 10 15  
 Ser Ala Thr Ser Ser Thr Leu Gly Thr Thr Gly Ser Val Pro Thr Ser  
 20 25 30  
 Thr Asp Pro Ala Pro Ser Ala His Leu Asp Ser Val His Lys Ser Thr  
 35 40 45  
 Asp Ser Gly Pro Ser Glu Leu Pro Gly Pro Thr His Thr Thr Thr Gly  
 50 55 60  
 Ser Thr Tyr Ser Ala Ile Thr Thr Thr His Ser Ala Pro Ser Pro Leu  
 65 70 75 80  
 Thr His Thr Thr Thr Gly Ser Thr His Lys Pro Ile Ile Ser Thr Leu  
 85 90 95

Thr Thr Thr Gly Pro Thr Leu Asn Ile Ile Gly Pro Val Gln Thr Thr  
 100 105 110  
 Thr Ser Pro Thr His Thr Met Pro Ser Pro Thr His Thr Thr Ala Ser  
 115 120 125  
 Pro Thr His Thr Ser Thr Ser Pro Thr His Thr Pro Thr Ser Pro Thr  
 130 135 140  
 His Lys Thr Ser Met Ser Pro Pro Thr Thr Thr Ser Pro Thr Pro Ser  
 145 150 155 160  
 Gly Met Gly Leu Val Gln Thr Ala Thr Ser Pro Thr His Pro Thr Thr  
 165 170 175  
 Ser Pro Thr His Pro Thr Thr Ser Pro Ile Leu Ile Asn Val Ser Pro  
 180 185 190  
 Ser Thr Ser Leu Glu Leu Ala Thr Leu Ser Ser Pro Ser Lys His Ser  
 195 200 205  
 Asp Pro Thr Leu Pro Gly Thr Asp Ser Leu Pro Cys Ser Pro Pro Val  
 210 215 220  
 Ser Asn Ser Tyr Thr Gln Ala Asp Pro Met Ala Pro Arg Thr Pro His  
 225 230 235 240  
 Pro Ser Pro Ala His Ser Ser Arg Lys Pro Leu Thr Ser Pro Ala Pro  
 245 250 255  
 Asp Pro Ser Glu Ser Thr Val Gln Ser Leu Ser Pro Thr Pro Ser Pro  
 260 265 270  
 Pro Thr Pro Ala Pro Gln His Ser Asp Leu Cys Leu Ala Met Ala Val  
 275 280 285  
 Gln Thr Pro Val Pro Thr Ala Ala Gly Gly Ser Gly Asp Arg Ser Leu  
 290 295 300  
 Glu Glu Ala Leu Gly Ala Leu Met Ala Ala Leu Asp Asp Tyr Arg Gly  
 305 310 315 320  
 Gln Phe Pro Glu Leu Gln Gly Leu Glu Gln Glu Val Thr Arg Leu Glu  
 325 330 335  
 Ser Leu Leu Met Gln Arg Gln Gly Leu Thr Arg Ser Arg Ala Ser Ser  
 340 345 350  
 Leu Ser Ile Thr Val Glu His Ala Leu Glu Ser Phe Ser Phe Leu Asn  
 355 360 365  
 Glu Asn Glu Asp Glu Asp Asn Asp Val Pro Gly Asp Arg Pro Pro Ser  
 370 375 380  
 Ser Pro Glu Ala Gly Ala Glu Asp Ser Ile Asp Ser Pro Ser Ala Arg

385                      390                      395                      400  
 Pro Leu Ser Thr Gly Cys Pro Ala Leu Asp Ala Ala Leu Val Arg His  
                                  405                      410                      415  
 Leu Tyr His Cys Ser Arg Leu Leu Leu Lys Leu Gly Thr Phe Gly Pro  
                                  420                      425                      430  
 Leu Arg Cys Gln Glu Ala Trp Ala Leu Glu Arg Leu Leu Arg Glu Ala  
                                  435                      440                      445  
 Arg Val Leu Glu Ala Val Cys Glu Phe Ser Arg Arg Trp Glu Ile Pro  
                                  450                      455                      460  
 Ala Ser Ser Ala Gln Glu Val Val Gln Phe Ser Ala Ser Arg Pro Gly  
                                  465                      470                      475                      480  
 Phe Leu Thr Phe Trp Asp Gln Cys Thr Glu Arg Leu Ser Cys Phe Leu  
                                  485                      490                      495  
 Cys Pro Val Glu Arg Val Leu Leu Thr Phe Cys Asn Gln Tyr Gly Ala  
                                  500                      505                      510  
 Arg Leu Ser Leu Arg Gln Pro Gly Leu Ala Glu Ala Val Cys Val Lys  
                                  515                      520                      525  
 Phe Leu Glu Asp Ala Leu Gly Gln Lys Leu Pro Arg Arg Pro Gln Pro  
                                  530                      535                      540  
 Gly Pro Gly Glu Gln Leu Thr Val Phe Gln Phe Trp Ser Phe Val Glu  
                                  545                      550                      555                      560  
 Thr Leu Asp Ser Pro Thr Met Glu Ala Tyr Val Thr Glu Thr Ala Glu  
                                  565                      570                      575  
 Glu Val Leu Leu Val Arg Asn Leu Asn Ser Asp Asp Gln Ala Val Val  
                                  580                      585                      590  
 Leu Lys Ala Leu Arg Leu Ala Pro Glu Gly Arg Leu Arg Arg Asp Gly  
                                  595                      600                      605  
 Leu Arg Ala Leu Ser Ser Leu Leu Val His Gly Asn Asn Lys Val Met  
                                  610                      615                      620  
 Ala Ala Val Ser Thr Gln Leu Arg Ser Leu Ser Leu Gly Pro Thr Phe  
                                  625                      630                      635                      640  
 Arg Glu Arg Ala Leu Leu Cys Phe Leu Asp Gln Leu Glu Asp Glu Asp  
                                  645                      650                      655  
 Val Gln Thr Arg Val Ala Gly Cys Leu Ala Leu Gly Cys Ile Lys Ala  
                                  660                      665                      670  
 Pro Glu Gly Ile Glu Pro Leu Val Tyr Leu Cys Gln Thr Asp Thr Glu  
                                  675                      680                      685

Ala Val Arg Glu Ala Ala Arg Gln Ser Leu Gln Gln Cys Gly Glu Glu  
690 695 700

Gly Gln Ser Ala His Arg Arg Leu Glu Glu Ser Leu Asp Ala Leu Pro  
705 710 715 720

Arg Ile Phe Gly Pro Gly Ser Met Ala Ser Thr Ala Phe  
725 730

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2501 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```

GGCAGTCAAA GCACTGATGA TGATAAAATC GTTCAGTACC ATTGGGAAGA ACTTAAGGGG      60
CCTCTAAGAG AAGAGAAGAT TTCTGAAGAT ACAGCCATAT TAAACTAAG TAAACTCGTC      120
CCTGGGAACT ACACTTTCAG CTTGACTGTA GTAGACTCTG ATGGAGCTAC CAACTCTACT      180
ACTGCAAACC TGACAGTGAA CAAAGCTGTG GATTACCCCC CTGTGGCCAA CGCAGGCCCC      240
AACCAAGTGA TCACCCTGCC CCAAACTCC ATCACCCTCT TTGGGAACCA GAGCACTGAT      300
GATCATGGCA TCACCAGCTA TGAGTGGTCA CTCAGCCCAA GCAGCAAAGG GAAAGTGGTG      360
GAGATGCAGG GTGTTAGAAC ACCAACCTTA CAGCTCTCTG CGATGCAAGA AGGAGACTAC      420
ACTTACCAGC TCACAGTGAC TGACACAATA GGACAGCAGG CCACTGCTCA AGTGACTGTT      480
ATTGTGCAAC CTGAAAACAA TAAGCCTCCT CAGGCAGATG CAGGCCCAGA TAAAGAGCTG      540
ACCCTTCCTG TGGATAGCAC AACCTGGAT GGCAGCAAGA GCTCAGATGA TCAGAAAATT      600
ATCTCATATC TCTGGGAAAA AACACAGGGA CCTGATGGGG TGCAGCTCGA GAATGCTAAC      660
AGCAGTGTTG CTACTGTGAC TGGGCTGCAA GTGGGGACCT ATGTGTTTAC CTTGACTGTC      720
AAAGATGAGA GGAACCTGCA AAGCCAGAGC TCTGTGAATG TCATTGTCAA AGAAGAATAA      780
ACAAACCACC TATAGCCAAG ATAAGTGGGA ATGTGGTGAT TACCCTACCC ACGAGCACAG      840
CAGAGCTGGA TGGCTCTAAG TCCTCAGATG ACAAGGGAAT AGTCAGCTAC CTCTGGACTC      900
GAGATGAGGG GAGCCCAGCA GCAGGGGAGG TGTAAATCA CTCTGACCAT CACCCTATCC      960

```

TTTTTCTTTC AAACCTGGTT GAGGGAACCT ACACTTTTCA CCTGAAAGTG ACCGATGCAA	1020
AGGGTGAGAG TGACACAGAC CGGACCACTG TGGAGGTGAA ACCTGATCCC AGGAAAAACA	1080
ACCTGGTGGG GATCATCTTG GATATCAACG TCAGTCAGCT AACTGAGAGG CTGAAGGGGA	1140
TGTTTCATCCG CCAGATTGGG GTCCTCCTGG GGGTGCTGGA TTCCGACATC ATTGTGCAA	1200
AGATTTCAGCC GTACACGGAG CAGAGCACCA AAATGGTATT TTTTGTTCAG AACGAGCCTC	1260
CCCACCAGAT CTTCAAAGGC CATGAGGTGG CAGCGATGCT CAAGAGTGAG CTGCGGAAGC	1320
AAAAGGCAGA CTTTTTGATA TTCAGAGCCT TGGAAAGTCAA CACTGTCACA TGTCAGCTGA	1380
ACTGTTCCGA CCATGGCTAC TGTGACTCGT TCACCAAACG CTGTATCTGT GACCCTTTTT	1440
GGATGGAGAA TTTCATCAAG GTGCAGCTGA GGGATGGAGA CAGCAACTGT GAGTGGAGCG	1500
TGTTATATGT TATCATTGCT ACCTTTGTCA TTGTTGTTGC CTTGGGAATC CTGTCTTGGA	1560
CTGTGATCTG TTGTTGTAAG AGGC AAAAG GAAAACCCAA GAGGAAAAGC AAGTACAAGA	1620
TCTGGATGCC ACGGATCAGG AAAGCTTGGA GCTGAAGCCA ACCTCCCGAG CAGGCATCAA	1680
ACAGAAAGGC CTTTTGCTAA GTAGCAGCCT GATGCACTCC GAGTCAGAGC TGGACAGCGA	1740
TGATGCCATC TTTACATGGC CAGACCGAGA GAAGGGCAAA CTCCTGCATG GTCAGAATGG	1800
CTCTGTACCC AACGGGCAGA CCCCTCTGAA GGCCAGGAGC CCGCGGGAGG AGATCCTGTA	1860
GCCACCTGGT CTGTCTCCTC AGGGCAGGGC CCAGCACACT GCGCGGCCAG TCCTCCTACC	1920
TCCCGAGTCT GCGGGCAGCT GCTGTCCCAG CATCTGCTGG TCATTTCGCC CTGACAGTCC	1980
CAACCAGAAC CCCTGGGACT TGAATCCAGA GACGTCTTTC AGGAACCCCT CAACGAAGCT	2040
GTGAATGAAG AGGTTTCCTC TTAAACCTG TCTGGTGGGC CCCCAGATAT CCTCACCTCA	2100
GGGCCTCCTT TTTTTCGAAA CTCCTCCCCT CCCCCGAGGG CAGACCCAGC CAGCTGCTAA	2160
GCTCTGCAGC TCCCCAGTGG ACAGTGTCTAT TGTGCCCAGA GTGCTGCAAG GTGAGGCCTG	2220
CTGTGCTGCC CGCACACCTG AGTGCAAAAC CAAGCACTGT GGGCATGGTG TTTCCCTCTC	2280
TGGGGTAGAG TACGCCCTCT CGCTGGGCAA AGAGGAAGTG GCACCCCTCC CCTCACCACA	2340
GATGCTGAGA TGGTAGCATA GAAATGATGG CCGGGCGCGG TGGCTMACGC CTGTAATCCC	2400
AGCACTTTGG GAGGCCGAGG CGGGCGGATC ATGAGGTCAG GAGATCAAGA CCACCCTGGC	2460
TAACACGGTG AAACCCCATC TCTACTAAAA AAAAAAAAAA A	2501

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 138 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

```

Met Gln Gly Val Arg Thr Pro Thr Leu Gln Leu Ser Ala Met Gln Glu
1           5           10           15
Gly Asp Tyr Thr Tyr Gln Leu Thr Val Thr Asp Thr Ile Gly Gln Gln
20           25           30
Ala Thr Ala Gln Val Thr Val Ile Val Gln Pro Glu Asn Asn Lys Pro
35           40           45
Pro Gln Ala Asp Ala Gly Pro Asp Lys Glu Leu Thr Leu Pro Val Asp
50           55           60
Ser Thr Thr Leu Asp Gly Ser Lys Ser Ser Asp Asp Gln Lys Ile Ile
65           70           75           80
Ser Tyr Leu Trp Glu Lys Thr Gln Gly Pro Asp Gly Val Gln Leu Glu
85           90           95
Asn Ala Asn Ser Ser Val Ala Thr Val Thr Gly Leu Gln Val Gly Thr
100          105          110
Tyr Val Phe Thr Leu Thr Val Lys Asp Glu Arg Asn Leu Gln Ser Gln
115          120          125
Ser Ser Val Asn Val Ile Val Lys Glu Glu
130          135

```

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1820 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGCAGCAGCA GAGGGAGAGC TCGGGGCTTG GAGGGGAAAC AGCGGAAGAC CTAAGATTAT

60



CGGGAGGGCA GCAGAGGCAG AGAACGAGGA CAGGACCCTT GGCCGTCTTC TTCCAGGGAA 120  
 CGAGAGGTCA CAGCCTCGCT CTCCGCTTAG GCTTCTGGCG CCCCAGCTTA AAGCCGAGGC 180  
 TGCGGCTGAC AAAGGGCTCG CGCCGGTGCC GCCGCCCTTC TCATCCGGGC ATTCCGGGTCC 240  
 CTGCGGAGAG GGAGGGGGAA GGGCAGAGGG GGAGGGGAAG GAGCCSGAGG GGCSCACACT 300  
 TGGAGCTGAA GCCCTCTCCA GGGCTCCGGG CCGGTGCCCC AACGGACAGA GGTTCGAGGAG 360  
 GACCCGCAGA GGTGGCAGCG GCCGGGGGCA GGAGGATGGT GCAGAAGGAG AGTCAAGCGA 420  
 CGTTGGAGGA GCGGGAGAGC GAGCTCAGTT CCAACCCTGC CGCCTCTGCG GGGGCATCGC 480  
 TGGAGCCGCC GGCAGCTCCG GCACCCGGAG AAGACAACCC CGCCGGGGCT GGGGGAGCGG 540  
 CCGTGGCCCG GGCTGCAGGA GGGGCTCGGC GGTTCCTCT GCGGTGTGGT GGAAGGATTT 600  
 TATGGAAGAC CTTGGGTTAT GGAACAGAGA AAAGAACTCT TTAGAAGGCT CCAGAAATGG 660  
 GAATTAAATA CATACTTGTA TGCCCCAAAA GATGACTACA AACATAGGAT GTTTTGGCGA 720  
 GAGATGTATT CAGTGGAGGA AGCTGAGCAA CTTATGACTC TCATCTCTGC TGCACGAGAA 780  
 TATGAGATAG AGTTCATCTA TGCGATCTCA CCTGGATTGG ATATCACTTT TTCTAACCCC 840  
 AAGGAAGTAT CCACATTGAA ACGTAAATTG GACCAGGTTT CTCAGTTTGG GTGCAGATCA 900  
 TTTGCTTTGC TTTTGTATGA TATAGACCAT AATATGTGTG CAGCAGACAA AGAGGTATTC 960  
 AGTTCTTTTG CTCATGCCCA AGTCTCCATC ACAAATGAAA TCTATCAGTA CCTAGGAGAG 1020  
 CCAGAAACTT TCCTCTTCTG TCCCACAGAA TACTGTGGCA CTTTCTGTTA TCCAAATGTG 1080  
 TCTCAGTCTC CATATTTAAG GACTGTGGGT GAAAAGCTTC TACCTGGAAT TGAAGTGCTT 1140  
 TGGACAGGTC CCAAAGTTGT TTCTAAAGAA ATTCCAGTAG AGTCCATCGA AGAGGTTTCT 1200  
 AAGATTATTA AGAGAGCTCC AGTAATCTGG GATAACATTC ATGCTAATGA TTATGATCAG 1260  
 AAGAGACTGT TTCTGGGCCC GTACAAAGGA AGATCCACAG AACTCATCCC ACGGTTAAAA 1320  
 GGAGTCCTCA CTAATCCAAA TTGTGAATTT GAAGCCAACT ACGTTGCTAT CCACACCCTT 1380  
 GCCACCTGGT ACAAATCAAA CATGAATGGG AGTGAGAAAA GATGTAGTGA TGAAGTACAG 1440  
 TGAAGATAGT ACTGTGTCCA TCCAGATAAA ATTAGAAAAT GAAGGCAGTG ATGAAGATAT 1500  
 TGAAACTGAT GTAAGTCTATA GTCCACAGAT GGCTCTAAAG CTAGCATTAA CAGAATGGTT 1560  
 GCAAGAGTTT GGTGTGCTC ATCAATACAG CAGTAGGCAA GTTGCACACA GTGGAGCTAA 1620  
 AGCAAGTGTA GTTGATGGGA CTCCTTTAGT TGCAGCACCC TCTTTAAATG CCACAACCGT 1680  
 AGTAACAACA GTTTATCAGG AGCCCATAT GAGCCAGGGA GCAGCCTTGA GTGGTGAGCC 1740

TACTACTCTG ACCAAGGAAG AAGAAAAGAA ACAGCCTGAT GAAGAACCCA TGGACATGGT 1800  
 GGTGGAAAAA AAAAAAAAAA 1820

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 272 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Glu Gln Arg Lys Glu Leu Phe Arg Arg Leu Gln Lys Trp Glu Leu  
 1 5 10 15  
 Asn Thr Tyr Leu Tyr Ala Pro Lys Asp Asp Tyr Lys His Arg Met Phe  
 20 25 30  
 Trp Arg Glu Met Tyr Ser Val Glu Glu Ala Glu Gln Leu Met Thr Leu  
 35 40 45  
 Ile Ser Ala Ala Arg Glu Tyr Glu Ile Glu Phe Ile Tyr Ala Ile Ser  
 50 55 60  
 Pro Gly Leu Asp Ile Thr Phe Ser Asn Pro Lys Glu Val Ser Thr Leu  
 65 70 75 80  
 Lys Arg Lys Leu Asp Gln Val Ser Gln Phe Gly Cys Arg Ser Phe Ala  
 85 90 95  
 Leu Leu Phe Asp Asp Ile Asp His Asn Met Cys Ala Ala Asp Lys Glu  
 100 105 110  
 Val Phe Ser Ser Phe Ala His Ala Gln Val Ser Ile Thr Asn Glu Ile  
 115 120 125  
 Tyr Gln Tyr Leu Gly Glu Pro Glu Thr Phe Leu Phe Cys Pro Thr Glu  
 130 135 140  
 Tyr Cys Gly Thr Phe Cys Tyr Pro Asn Val Ser Gln Ser Pro Tyr Leu  
 145 150 155 160  
 Arg Thr Val Gly Glu Lys Leu Leu Pro Gly Ile Glu Val Leu Trp Thr  
 165 170 175  
 Gly Pro Lys Val Val Ser Lys Glu Ile Pro Val Glu Ser Ile Glu Glu  
 180 185 190

Val Ser Lys Ile Ile Lys Arg Ala Pro Val Ile Trp Asp Asn Ile His  
 195 200 205

Ala Asn Asp Tyr Asp Gln Lys Arg Leu Phe Leu Gly Pro Tyr Lys Gly  
 210 215 220

Arg Ser Thr Glu Leu Ile Pro Arg Leu Lys Gly Val Leu Thr Asn Pro  
 225 230 235 240

Asn Cys Glu Phe Glu Ala Asn Tyr Val Ala Ile His Thr Leu Ala Thr  
 245 250 255

Trp Tyr Lys Ser Asn Met Asn Gly Ser Glu Lys Arg Cys Ser Asp Asp  
 260 265 270

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2405 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19::

TATCCATTAC GTCGACTAAT ACGTACATAA GAATTCAATC GGGAGACAAT GTAATAACCC 60

AAACACTGGG TATTCATATT TGACACATGG GCAAACCTGC CAGTGAATG GAATTGTGAC 120

CTGACAGAGA AGGGAAGGCA GGCTGACGAA GGTGATCGAA TGGGAGAACA GTGCTGTGGT 180

GATCATGAGA ATGAGGCTTT TCTGTAGCAT GTAAACCAA CCGGACCCTT GGCAGTTCGT 240

CGTCCCTCAG TTCTCCAGAT GCTATTTTTT GCAGGTTCTA CCAAGTGCTT GTTGATTACC 300

CTAGTTGTAA TTATCTAGGG AAGAGATGAA TGTAAGTGAG AGTGCAGAGC ACTGGGGAGG 360

GTGACAGTGA AATGCAATTA GAGGCAGCAA GAGAGTCCTA GTCTGTTCTC ACATAACCGG 420

ACTTGAACGC TCCAGTGCGA GCAGAGTGCT GGGGGTGGAT TCCACTGCCG AACCACGGCA 480

GCTTTGCTTT ACTCTTCAGC ATGGGGGTGG TAACTAGCTG CACAGCAAGT TATGAAATGG 540

AAAGCAAGCT TAACAGCTGT AATCTCATCG GATACCCTGG AGCAAATGCC TTGGGATTGC 600

CTGAAGTGAA GTGTTTAGCA TCCACCAAAT AGTTGAGTTT CTAAGATGGG CCATGCGGGA 660

TCCCTGCCAC ACGGGTGTGG GGGCAGCGCG CTCCTCCCTG CCTCGGCTGC TGTGTGTCTT 720

AGTCTTCATG TTTCCCCTTG GCGTCACCTC TGGTCCTGTC CATCTTGTGC ACATGTGTCC	780
CTGAGGGCGA GTCACACCTG TGTCTTGAGT CTGCTGTAAG GTGTTCAGTC TACCTCAAGG	840
GGTCATCATG GTAAGCTCTA TCCCACGGCA TCCTCCGTCC ATCACCACAC AACTGGAAT	900
GTATCCCCC CGCCCTGCCC CTTCGATAAC CTCCTGACCT TGTCTTTCCT ATAATCACCA	960
CAAACTCCTA GTGCAAATTG GATGCTGCTT TAAATGTGAA AACAATTGTT CAAAAGCTAT	1020
AAAACCTGAA TGAAAGCTGA AGCTGAATTT ATAAGCCTTG TTGCATATGA GCCAAAAGTG	1080
CAAAAAGCTC TATATAATGA ACTAACCTGC CACTCGTATA AATATAAATA TATATAAATA	1140
TATTAAAATC AGACTGTTCT ACAAATTTGT GTATTGTAT TTTTGTGTAC GTACAATTTT	1200
CTGCTAAGCA GAAGGAGGAT GTAGTATAGG ATGCTGTGAG AAGAGATTTG GTTTAATCCT	1260
ATTTCTTTGT TACTTATTTT TGTTAACATC AGATGCCTGT CTTTGTCTTA TATGTGTATG	1320
AACTGTTTG GAAAGCTGCA GTATCTCTCT TCCTGAGGTC CAGAGGCCAT TAAAGAATAA	1380
ACATGGGGCT TAGAGGTGTC TTGGCCACCA AACACAGGTA AGCCAACAGA GTTCGGCTCC	1440
CGTACGTGGT GTGGTGGACC TGGGGCCTGG AGCCACTTGT ACTCCTGGGG TGGGACTGAG	1500
GGAGGGAAAG GGCTCCCCCA CTTCAGGCTG GTCACTGTAT GTTGAATTCC CTCTGGGAAG	1560
CAAAATTCCA CTTGTAAAAA TCAGCTTCCA AGACAAAAGA TGATCCGATG TCACTTCTGG	1620
GTCACCAGGA AGGGACACAG GATCTCTCTG AACTTTGGAA CAGACCCTCA TATTCTGGGG	1680
CCAGAAGTTT GCCCAGAAAG CAGCAGGTGG CTCTGCCTGG CTGTAGAGCC CAGCTCATTG	1740
GCTGTCCCTG GGCTTGGTCT CCCTCTTCCG AGTAGTTGCT GCCTTTCTTC AGATCAGGTT	1800
ACCACAATGC CTCCCCGCTG CTGACGCTTC ATCCCCACA CCTCCAGCCC CAGTTACCTG	1860
GAGCTTCTCA GAACCCACTT TGCCGGTGCT AAAACACAAG AGGGGGTGAA AGTGGCTGCC	1920
AGTAATGGCC AGAAACCAAC CACCAGAGGC CAGGCTGAAA GACAAGCTCC GGGTGTCCAG	1980
GGGCTGACGG GCCAACCATG TGGCAGGTCC CAGGCCCCAC CCACTGCGCC ATCCGCCTCT	2040
GAGCTCCACA GTGGTCCCAC TAATGGGAAC CTCCTCTAGG GAGAGTGATA CTGCACCTTC	2100
ACCCGTAGGA CTCATATTTA TAACAATGTG TAATGGCTGT AGCAAAAAGC CCTTGTCTCT	2160
AGATGTAAAT GGTCAAAGAA ACAAGCGCTC TATTGTTTTG AATAAAATAG TTCAAATGAG	2220
TCCTGTATCA TTGTATCTCC TATTCTGGAT TAGTGCCTTT TGGACAGTAG ACTGTTCTGT	2280
AATTAAAATG TAGTATACTG CTTTTTTGTA CAGTTTTGTT TTAATAAAAC TTTTTTTTAA	2340
TTTGTGTTTA TTTTAGTATT GTACCTATTA GAGAATAAAA TGTATAACTG AAAAAAAAAA	2400

AAAAA

2405

## (2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 76 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Gly Thr Ser Ser Arg Glu Ser Asp Thr Ala Pro Ser Pro Val Gly
1              5              10              15
Leu Ile Phe Ile Thr Met Cys Asn Gly Cys Ser Lys Lys Pro Leu Phe
20              25              30
Leu Asp Val Asn Gly Gln Arg Asn Lys Arg Ser Ile Val Leu Asn Lys
35              40              45
Ile Val Gln Met Ser Pro Val Ser Leu Tyr Leu Leu Phe Trp Ile Ser
50              55              60
Ala Phe Trp Thr Val Asp Cys Ser Val Ile Lys Met
65              70              75

```

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CNAGCAGCAGA TACAGCAGTA AGGAAGGC

29

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GNGCTCTCCTT AATCGCCGTC TCAAACAT

29

- (2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CNTTGGCCAGG TCTTCCTGCT CCTTCATT

29

- (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
  - (A) DESCRIPTION: /desc = "oligonucleotide"

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

ANGCTGCAAAG CTGGTGTAGA AGTAGGTG

29

- (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 29 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ANTTAGCCATG TCACTCCGCA GGAAGTCC

29

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TNTAGTCATAA AGAGCCACTA CCGTTGGT

29

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TNGTATGGGTA GGGCTTGGCA TAGTGTGG

29

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TNTCAGGTTTG CAGTAGTAGA GTTGGTAG

29

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CNTCTTCGATG GACTCTACTG GAATTTCT

29

(2) INFORMATION FOR SEQ ID NO:30:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 29 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid  
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

TNAGCCTGGCC TCTGGTGGTT GGTTTCTG

29

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 336 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein



(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```

Thr Leu Lys Leu Cys Ser Pro Pro Lys Asp His Glu Val Leu Gly Val
1           5           10           15
Ile Gln Arg Phe Leu Lys Leu Arg Ser Pro Trp Lys Ala Leu Ile Thr
20           25           30
Thr Ile Leu Lys Pro Gln Leu Phe Glu Pro Arg Pro Arg Leu Thr Val
35           40           45
Leu Thr Phe Pro Ser Ser His Glu Gly Glu Ser Phe Pro Leu Ala Trp
50           55           60
Asn Ala Lys Ile Thr Asp Leu Lys Gln Lys Val Glu Asn Leu Phe Asn
65           70           75           80
Glu Lys Cys Gly Glu Ala Leu Gly Leu Lys Gln Ala Val Lys Val Pro
85           90           95
Phe Ala Leu Phe Glu Ser Phe Pro Glu Asp Phe Tyr Val Glu Gly Leu
100          105          110
Pro Glu Gly Val Pro Phe Arg Arg Pro Ser Thr Phe Gly Ile Pro Arg
115          120          125
Leu Glu Lys Ile Leu Arg Asn Lys Ala Lys Ile Lys Phe Ile Ile Lys
130          135          140
Lys Pro Glu Met Phe Glu Thr Ala Ile Lys Glu Ser Thr Ser Ser Lys
145          150          155          160
Ser Pro Pro Arg Lys Ile Asn Ser Ser Pro Asn Val Asn Thr Thr Ala
165          170          175
Ser Gly Val Glu Asp Leu Asn Ile Ile Gln Val Thr Ile Pro Asp Asp
180          185          190
Asp Asn Glu Arg Leu Ser Lys Val Glu Lys Ala Arg Gln Leu Arg Glu
195          200          205
Gln Val Asn Asp Leu Phe Ser Arg Lys Phe Gly Glu Ala Ile Gly Met
210          215          220
Gly Phe Pro Val Lys Val Pro Tyr Arg Lys Ile Thr Ile Asn Pro Gly
225          230          235          240
Cys Val Val Val Asp Gly Met Pro Pro Gly Val Ser Phe Lys Ala Pro
245          250          255
Ser Tyr Leu Glu Ile Ser Ser Met Arg Arg Ile Leu Asp Ser Ala Glu
260          265          270

```

Phe Ile Lys Phe Thr Val Ile Arg Pro Phe Pro Gly Leu Val Ile Asn  
 275 280 285

Asn Gln Leu Val Asp Gln Ser Glu Ser Lys Gly Pro Val Ile Gln Glu  
 290 295 300

Ser Ala Glu Pro Ser Gln Leu Glu Val Pro Ala Thr Glu Glu Ile Lys  
 305 310 315 320

Glu Thr Asp Gly Ser Ser Gln Ile Lys Gln Glu Pro Asp Pro Thr Trp  
 325 330 335

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

ANCTTGGCCA GGTCTTCCTG CTCCTTCAT

29

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 521 to nucleotide 1111;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 536 to nucleotide 817;
  - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ax318\_3 deposited under accession number ATCC 98353;
  - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353;
  - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ax318\_3 deposited under accession number ATCC 98353;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:2;
  - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
  - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
  - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.
4. The host cell of claim 3, wherein said cell is a mammalian cell.
5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:
  - (a) growing a culture of the host cell of claim 3 in a suitable culture medium; and
  - (b) purifying said protein from the culture.
6. A protein produced according to the process of claim 5.
7. The protein of claim 6 comprising a mature protein.
8. A protein comprising an amino acid sequence selected from the group consisting of:
  - (a) the amino acid sequence of SEQ ID NO:2;
  - (b) the amino acid sequence of SEQ ID NO:2 from amino acid 4 to amino acid 99;
  - (c) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 93 to amino acid 102 of SEQ ID NO:2; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone ax318\_3 deposited under accession number ATCC 98353;the protein being substantially free from other mammalian proteins.
9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.
10. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2 from amino acid 4 to amino acid 99.
11. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

12. A method for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically effective amount of a composition of claim 11.

13. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

14. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 61 to nucleotide 864;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 826 to nucleotide 1386;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;
- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bg140\_1 deposited under accession number ATCC 98353;
- (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;
- (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;
- (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:4;
- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

15. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:4;
  - (b) the amino acid sequence of SEQ ID NO:31;
  - (c) the amino acid sequence of SEQ ID NO:31 from amino acid 148 to amino acid 249;
  - (d) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:4;
  - (e) fragments of the amino acid sequence of SEQ ID NO:31 comprising the amino acid sequence from amino acid 163 to amino acid 172 of SEQ ID NO:31; and
  - (f) the amino acid sequence encoded by the cDNA insert of clone bg140\_1 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.

16. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

17. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 77 to nucleotide 1624;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 390 to nucleotide 789;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bg465\_2 deposited under accession number ATCC 98353;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353;

- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bg465\_2 deposited under accession number ATCC 98353;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 253 to amino acid 262 of SEQ ID NO:6;
  - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
  - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
  - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
18. A protein comprising an amino acid sequence selected from the group consisting of:
- (a) the amino acid sequence of SEQ ID NO:6;
  - (b) the amino acid sequence of SEQ ID NO:6 from amino acid 260 to amino acid 343;
  - (c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 253 to amino acid 262 of SEQ ID NO:6; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone bg465\_2 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.
19. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.
20. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 48 to nucleotide 1055;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 216 to nucleotide 1055;
  - (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 494 to nucleotide 958;
  - (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353;
  - (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bk291\_3 deposited under accession number ATCC 98353;
  - (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bk291\_3 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bk291\_3 deposited under accession number ATCC 98353;
  - (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;
  - (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:8;
  - (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
  - (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
  - (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).
21. A protein comprising an amino acid sequence selected from the group consisting of:
- (a) the amino acid sequence of SEQ ID NO:8;



- (b) the amino acid sequence of SEQ ID NO:8 from amino acid 188 to amino acid 306;
  - (c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 85 to amino acid 94 of SEQ ID NO:8; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone bk291\_3 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.

22. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.
23. An isolated polynucleotide selected from the group consisting of:
- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 64 to nucleotide 1197;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 1 to nucleotide 828;
  - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353;
  - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353;
  - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone bp537\_4 deposited under accession number ATCC 98353;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 184 to amino acid 193 of SEQ ID NO:10;

- (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
- (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
- (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

24. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:10;
  - (b) the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 255;
  - (c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 184 to amino acid 193 of SEQ ID NO:10; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone bp537\_4 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.

25. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

26. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 581 to nucleotide 1534;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 928 to nucleotide 1510;
- (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353;
- (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;

- (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cs431\_2 deposited under accession number ATCC 98353;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 154 to amino acid 163 of SEQ ID NO:12;
  - (j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;
  - (k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and
  - (l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).
27. A protein comprising an amino acid sequence selected from the group consisting of:
- (a) the amino acid sequence of SEQ ID NO:12;
  - (b) the amino acid sequence of SEQ ID NO:12 from amino acid 150 to amino acid 310;
  - (c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 154 to amino acid 163 of SEQ ID NO:12; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone cs431\_2 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.
28. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.
29. An isolated polynucleotide selected from the group consisting of:

- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;
- (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 29 to nucleotide 2227;
- (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1334 to nucleotide 2227;
- (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1 to nucleotide 746;
- (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353;
- (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;
- (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw976\_1 deposited under accession number ATCC 98353;
- (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;
- (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;
- (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 361 to amino acid 370 of SEQ ID NO:14;
- (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
- (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
- (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).

30. A protein comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence of SEQ ID NO:14;

- (b) the amino acid sequence of SEQ ID NO:14 from amino acid 1 to amino acid 239;
  - (c) the amino acid sequence of SEQ ID NO:14 from amino acid 119 to amino acid 733;
  - (d) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 361 to amino acid 370 of SEQ ID NO:14; and
  - (e) the amino acid sequence encoded by the cDNA insert of clone cw976\_1 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.

- 31. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.
- 32. An isolated polynucleotide selected from the group consisting of:
  - (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 364 to nucleotide 777;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 1 to nucleotide 636;
  - (d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353;
  - (e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;
  - (f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone cw1233\_3 deposited under accession number ATCC 98353;
  - (g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;
  - (i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment

comprising the amino acid sequence from amino acid 64 to amino acid 73 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

33. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 91;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 64 to amino acid 73 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone cw1233\_3 deposited under accession number ATCC 98353;  
the protein being substantially free from other mammalian proteins.

34. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

35. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 619 to nucleotide 1434;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:17 from nucleotide 520 to nucleotide 1323;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;

(f) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone dg1\_1 deposited under accession number ATCC 98353;

(g) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:18;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:18 having biological activity, the fragment comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:18;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(i).

36. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:18;

(b) the amino acid sequence of SEQ ID NO:18 from amino acid 1 to amino acid 235;

(c) fragments of the amino acid sequence of SEQ ID NO:18 comprising the amino acid sequence from amino acid 131 to amino acid 140 of SEQ ID NO:18; and

(d) the amino acid sequence encoded by the cDNA insert of clone dg1\_1 deposited under accession number ATCC 98353;  
the protein being substantially free from other mammalian proteins.

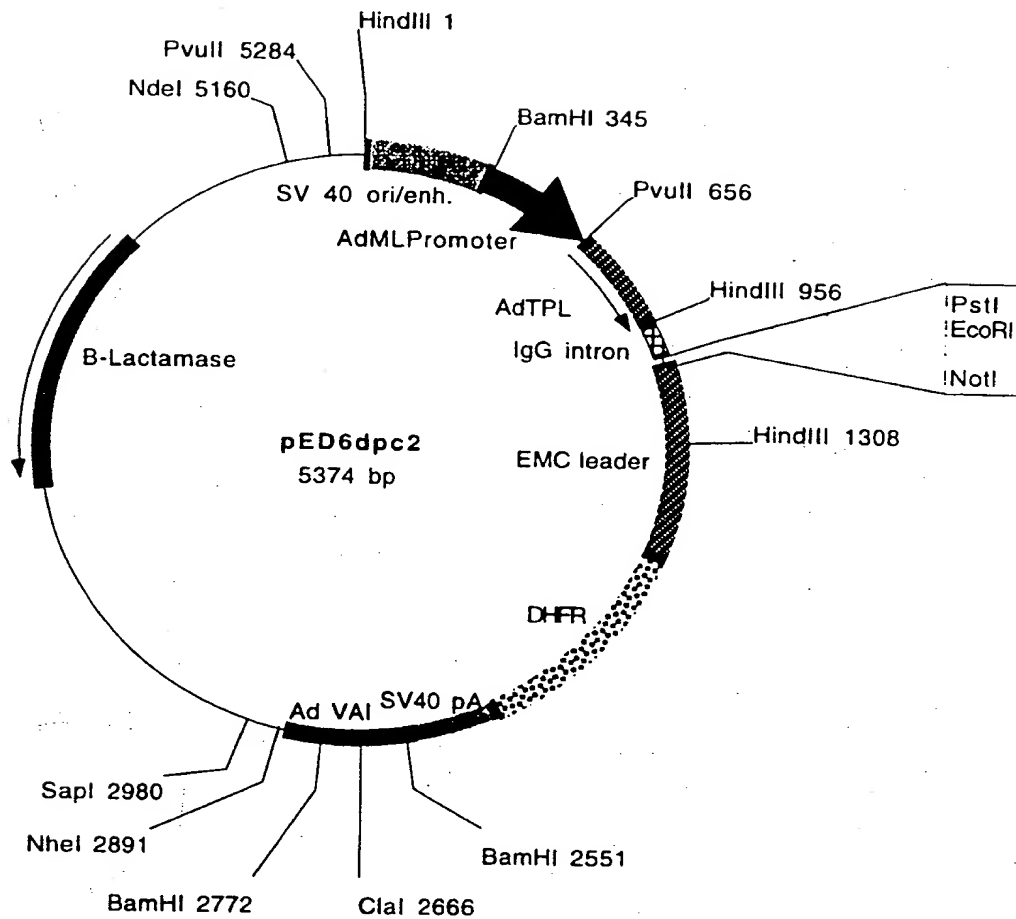
37. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:17.

38. An isolated polynucleotide selected from the group consisting of:
- (a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19;
  - (b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2063 to nucleotide 2290;
  - (c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2276 to nucleotide 2290;
  - (d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:19 from nucleotide 2037 to nucleotide 2405;
  - (e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353;
  - (f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;
  - (g) a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of clone ep234\_1 deposited under accession number ATCC 98353;
  - (h) a polynucleotide encoding the mature protein encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;
  - (i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:20;
  - (j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:20 having biological activity, the fragment comprising the amino acid sequence from amino acid 33 to amino acid 42 of SEQ ID NO:20;
  - (k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;
  - (l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and
  - (m) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(j).
39. A protein comprising an amino acid sequence selected from the group consisting of:



- (a) the amino acid sequence of SEQ ID NO:20;
  - (b) the amino acid sequence of SEQ ID NO:20 from amino acid 1 to amino acid 69;
  - (c) fragments of the amino acid sequence of SEQ ID NO:20 comprising the amino acid sequence from amino acid 33 to amino acid 42 of SEQ ID NO:20; and
  - (d) the amino acid sequence encoded by the cDNA insert of clone ep234\_1 deposited under accession number ATCC 98353;
- the protein being substantially free from other mammalian proteins.
40. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:19.

FIGURE 1A

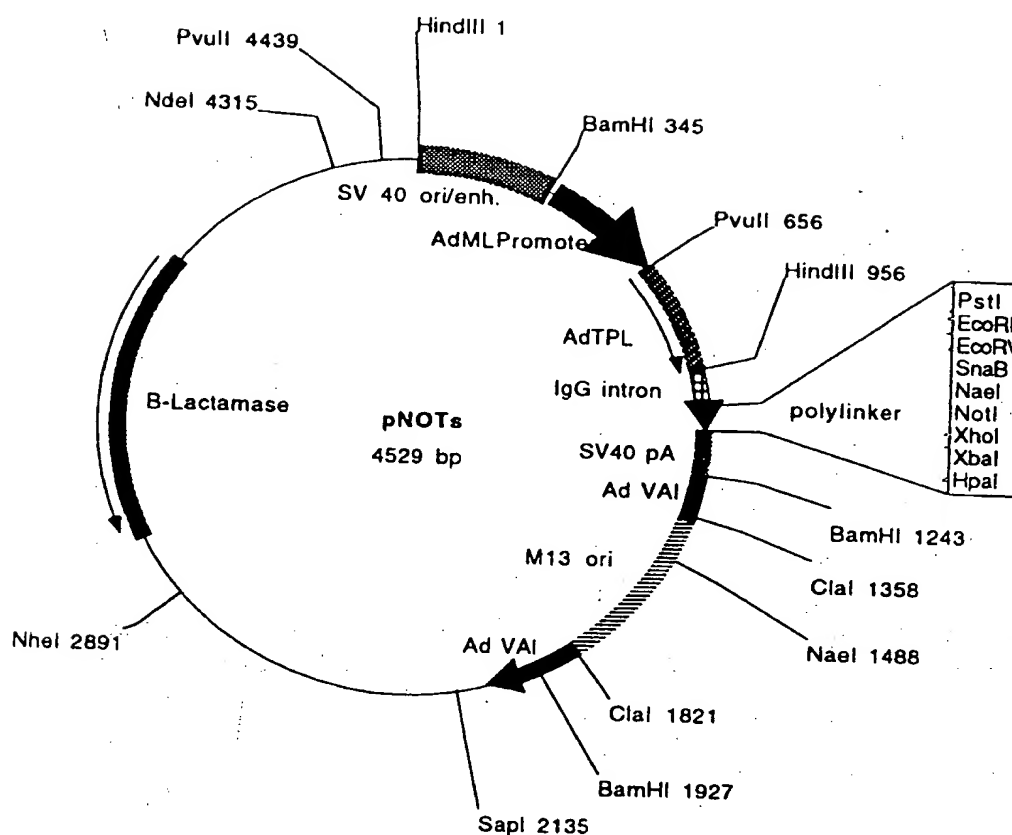


**Plasmid name:** pED6dpc2

**Plasmid size:** 5374 bp

**Comments/References:** pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs  
 Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al, 1989. Mol. Cell. Biol. 9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

**THIS PAGE BLANK (USPTO)**



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification<sup>6</sup>:

C12N 15/12, C07K 14/47, A61K 38/17

A3

(11) International Publication Number:

WO 98/40404

(43) International Publication Date: 17 September 1998 (17.09.98)

(21) International Application Number: PCT/US98/04601

(22) International Filing Date: 9 March 1998 (09.03.98)

(30) Priority Data:

08/815,381

11 March 1997 (11.03.97)

US

09/036,321

6 March 1998 (06.03.98)

US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 90 Green Meadow Drive, Tewksbury, MA 01876 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

## Published

*With international search report.**Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

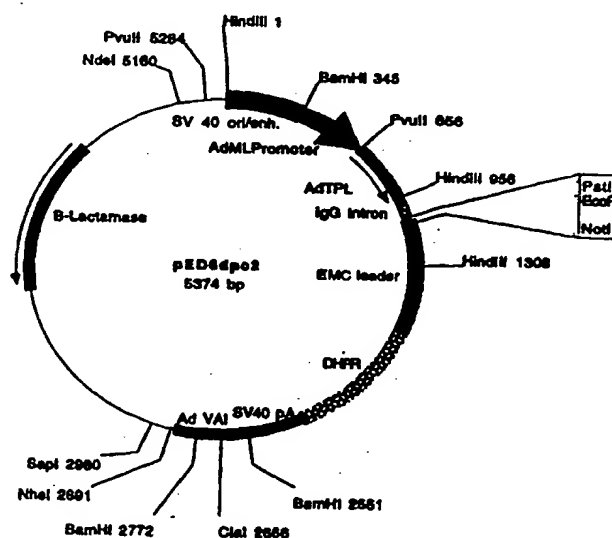
(88) Date of publication of the international search report:

30 December 1998 (30.12.98)

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

## (57) Abstract

Novel polynucleotides and the proteins encoded thereby are disclosed.



Plasmid name: pED6dpc2  
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. 88T cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al. (1991), NAR 19: 4485-4490.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## INTERNATIONAL SEARCH REPORT

International Application No.

PUS 98/04601

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C07K14/47 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 97 07198 A (GENETICS INST) 27 February 1997 see the whole document	1-11
A	US 5 536 637 A (K. JACOBS) 16 July 1996 cited in the application see the whole document	1-11
	--- -/-- ---	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*&\* document member of the same patent family

Date of the actual completion of the international search

4 August 1998

Date of mailing of the international search report

03.11.1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Mateo Rosell, A.M.

## INTERNATIONAL SEARCH REPORT

Inter. Application No  
PC 98/04601

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Y. FUJIOKA ET AL.,: "A novel membrane glycoprotein, SHPS-1, that binds the SH2-domain-containing protein tyrosine phosphatase SHP-2 in response to mitogens and cell adhesion" MOLECULAR AND CELLULAR BIOLOGY, vol. 16, no. 12, 1996, WASHINGTON, DC, US, pages 6887-6899, XP002072087 cited in the application see the whole document and specially Figure 4	1,13
P,X	--- DATABASE NUCLEOTIDE PROTEIN SEQUENCE, - 28 October 1997 HINXTON, GB, XP002072088 AC= AA625610 similar to WP:K10B2.1 CE02008 transducin beta-chain	1,11
P,X	--- DATABASE NUCLEOTIDE PROTEIN SEQUENCE, - 6 February 1998 HINXTON, GB, XP002072072 AC= AA778577 similar to human TR:P78324 SHPS-1	1,11
T	--- DATABASE NUCLEOTIDE PROTEIN SEQUENCE, - 15 March 1998 HINXTON, GB, XP002072089 AC= AA861049 similar to human TR:P78324 SHPS-1	1,11
T	--- DATABASE NUCLEOTIDE PROTEIN SEQUENCE, - 15 March 1998 XP002072090 AC= AA862667 similar to human TR:P78324 SHPS-1 -----	1,11



# INTERNATIONAL SEARCH REPORT

Ir 15 Application No.

PCT/US 98/ 04601

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
Remark: Although claim 12 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

## 1. Claims: 1-13

Polynucleotide comprising the nucleotide sequence of SEQ.ID.N.1 and encoding a polypeptide of SEQ.ID.N.2.  
Polynucleotide fragments, variants, homologues and gene thereof. Host cells transformed with said polynucleotide.  
Protein comprising SEQ.ID.N.2 or fragments thereof.  
Pharmaceutical compositions.

## 2. Claims: 14-16

As invention 1 but concerning SEQ.ID.N.3, 4 and 31

## 3. Claims: 17-19

As invention 1 but concerning SEQ.ID.N.5 and 6

## 4. Claims: 20-22

As invention 1 but concerning SEQ.ID.N.7 and 8

## 5. Claims: 23-25

As invention 1 but concerning SEQ.ID.N.9 and 10

## 6. Claims: 26-28

As invention 1 but concerning SEQ.ID.N.11 and 12

## 7. Claims: 29-31

As invention 1 but concerning SEQ.ID.N.13 and 14

## 8. Claims: 32-34

As invention 1 but concerning SEQ.ID.N.15 and 16

## 9. Claims: 35-37

As invention 1 but concerning SEQ.ID.N.17 and 18

## 10. Claims: 38-40

As invention 1 but concerning SEQ.ID.N.19 and 20

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. Application No

PCT/US 98/04601

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9707198	A	27-02-1997	US 5707829 A	13-01-1998
			AU 6712396 A	18-02-1997
			AU 6768596 A	12-03-1997
			EP 0839196 A	06-05-1998
			EP 0851875 A	08-07-1998
			WO 9704097 A	06-02-1997
US 5536637	A	16-07-1996	US 5712116 A	27-01-1998

**THIS PAGE BLANK (USPTO)**